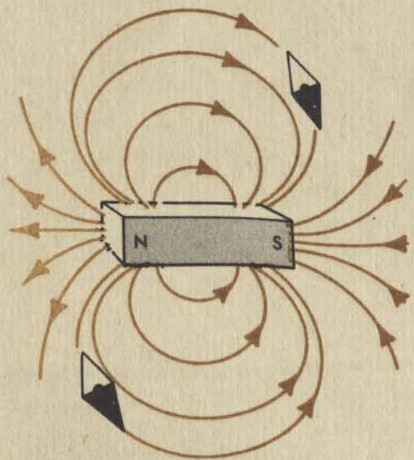
PHYSICS

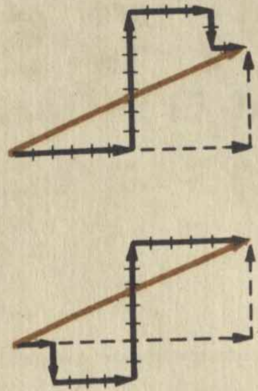
FOUNDATIONS AND FRONTIERS

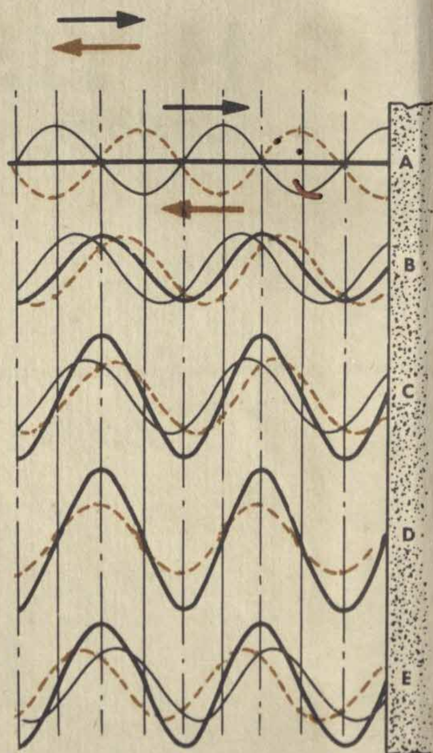
2nd EDITION

GEORGE GAMOW AND JOHN M. CLEVELAND

PHYSICS
Foundations and Frontiers


N
S




A
B
C
D
E

PHYSICS
Foundations and Frontiers

SECOND EDITION

George Gamow
Department of Physics
University of Colorado

John M. Cleveland
Department of Physics
University of Colorado

PRENTICE-HALL OF INDIA PRIVATE LIMITED
New Delhi, 1971

This Fourth Indian Reprint — Rs. 12.00
(Original U. S. Edition—Rs. 89.63)

PHYSICS : Foundations and Frontiers, 2nd Ed.,
by George Gamow and John M. Cleveland

PRENTICE-HALL INTERNATIONAL, INC., Englewood Cliffs.
PRENTICE-HALL OF INDIA PRIVATE LIMITED, New Delhi.
PRENTICE-HALL INTERNATIONAL INC., London.
PRENTICE-HALL OF CANADA LTD., Toronto.
PRENTICE-HALL OF JAPAN, INC., Tokyo.

This Eastern Economy Edition is the only authorised, complete and unabridged photo-offset reproduction of the latest American edition specially published and priced for sale only in Argentina, Bolivia, Brazil, Burma, Cambodia, Ceylon, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Ethiopia, Guatemala, Guyana, Haiti, Honduras, Hong Kong, India, Indonesia, Israel, Jamaica, Laos, Malaysia, Nicargua, Okinawa, Pakistan, Paraguay, Peru, Philippines, Singapore, South Korea, South Vietnam, Sudan, Surinam, Thailand, Tobago, Trinidad, Turkey, United Arab Republic, Uruguay and Venezuela.

Reprinted in India by special arrangement with Prentice-Hall, Inc., Englewood Cliffs, N. J., U.S.A.

First Printing (First Edition)	...	...	October, 1963
Second Printing (First Edition)	...	...	December, 1965
Third Printing (First Edition)	...	...	September, 1968
Fourth Printing (Second Edition)	...	...	August, 1971

This book has been published with the assistance of the joint Indian American Textbook Programme.

Printed by Mohan Makhijani at the Rekha Printers, New Delhi and Published by Prentice-Hall of India Private Limited, M-97, Connaught Circus, New Delhi-1.

Preface

Although the first edition of this book was in many ways an excellent text, our own experience in teaching from it, as well as the experience of other users, indicated that there were some areas in which revision was needed.

The chapter headings in the table of contents appear nearly the same as in the first edition—but within this framework a substantial amount of material has been rearranged to provide a superior integration and a more effectively teachable sequence of topics. Deletions and rewriting have been extensive, and much material has been added where appropriate to expand explanations and to bring our coverage up to date. In order to provide better support for student study, we have included many more worked-out examples.

A balance between deletions and additions has brought about the almost unique result that the amount of textual material in the second edition is no greater than that of the first. The book itself *is* longer, however, largely as a result of providing nearly three times as many

student questions and problems. We have included many more simple introductory questions, arranged in a sequence to give an almost programmed approach to the mastery of the material in each section. Marginal numbers indicate to the instructor the section to which each group of questions is pertinent—an aid, we hope, in making assignments. With few exceptions, the questions are in matched pairs: the odd-numbered question to which the student can find the answer in the back of the book, followed by the even-numbered question illustrating the same point, and to which the answer is *not* given. We see many pedagogic advantages to this scheme.

The preface to a book is written long after the more substantive material has been completed and delivered into the capable hands of the editors, illustrators, and compositors. It was during this hiatus that Dr. George Gamow died, leaving the entire scientific community with a sense of great loss, emphasized in my case by the realization that only one of us will be able to sign this preface. I hope that this book may be worthy of its place in the towering structure of words and ideas that constitute his enduring memorial.

JOHN M. CLEVELAND

3473

Contents

three

BODIES IN MOTION 38

four

ENERGY AND MOMENTUM 60

five

ROTATIONAL MECHANICS 81

six

ORBITS AND SATELLITES 95

eighteen

THE SPECIAL THEORY OF RELATIVITY 311

nineteen

THE GENERAL THEORY OF RELATIVITY 338

twenty

THE MOLECULAR NATURE OF MATTER 350

twenty-one

THE ELECTRICAL NATURE OF MATTER 370

chapter / one

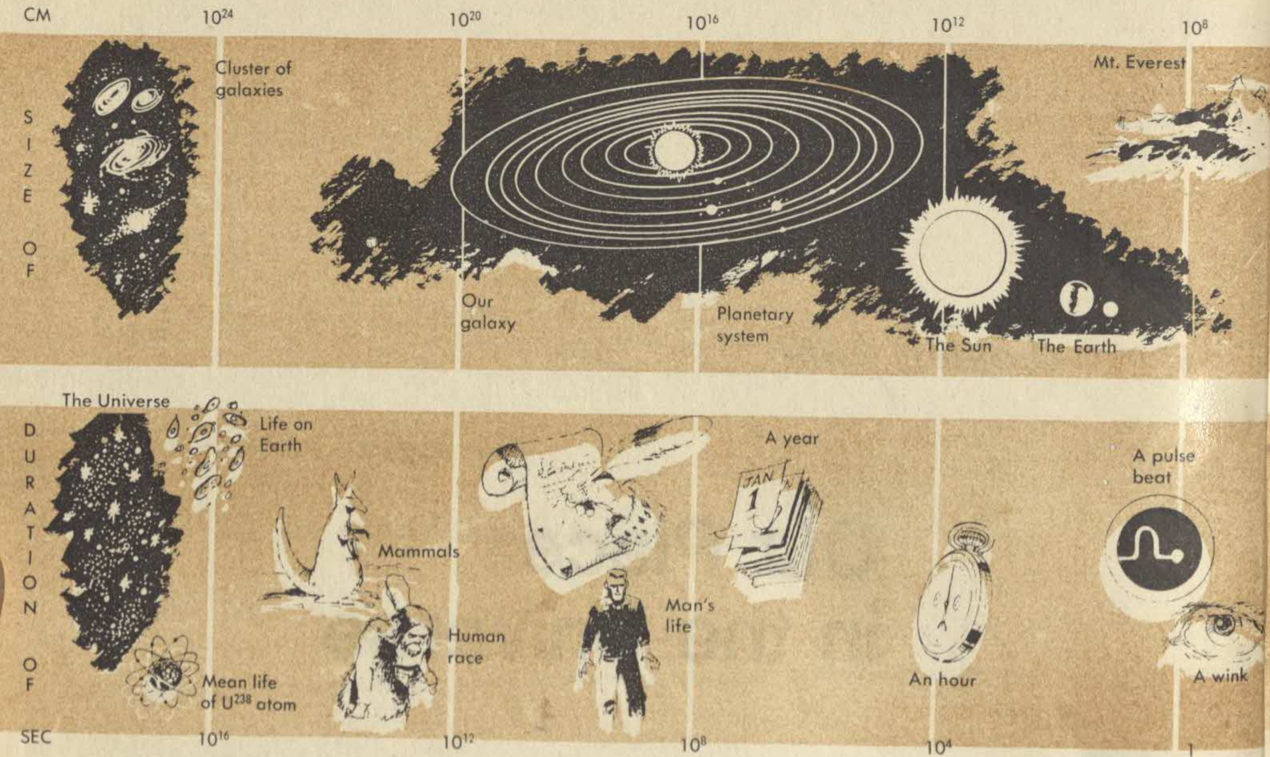
Our Place in the Universe

1-1 The Large and the Small

In our everyday life we encounter objects of widely differing sizes. Some of them are as large as a mountain and others as small as a grain of dust. When we go much beyond these limits, either in the direction of much larger objects or in the direction of much smaller ones, it becomes increasingly difficult to grasp their actual sizes.

Objects that are much larger than mountains, such as our earth itself, the moon, the sun, the stars, and stellar systems, constitute what is known as the *macrocosm* (Greek for "large world"). Very small objects such as bacteria, atoms, and electrons belong to the *microcosm* (Greek for "small world"). If we use one of the standard scientific units, the *meter* (39.37 inches) or the *centimeter* (0.01 meter; 2.54 centimeters = 1 inch) for measuring sizes, objects belonging to the macrocosm will be described by very large numbers, and those describing the microcosm by very small ones. Thus the diameter of the sun is 1,390,000,000 meters, and the diameter of a hydrogen atom is 0.0000000106 centimeter.

Scientists, however, would ordinarily express such very large and very small numbers in a different and more useful way. The diameter of the sun in meters is given by the number 139 followed by (and here we



must stop and count them) seven zeros. Since each zero means multiplication by 10, this corresponds to 139×10^7, or 1.39×10^9 meters. We can similarly express very small numbers in exponential notation if we recall that 10^{-3}, for example, means $1/10^3$, or 0.001. The diameter of the hydrogen atom is thus 1.06×10^{-8} centimeter.

To see how this "exponential notation" works, let us quickly review some of the rules for working with exponents. Suppose we want to multiply 10^2 by 10^5. Since $10^2 = 10 \times 10 = 100$, and $10^5 = 10 \times 10 \times 10 \times 10 \times 10 = 100{,}000$, these two numbers multiplied together give us 10,000,000, which is 10^7. It is easier, however, just to write $10^2 \times 10^5 = 10^7$; thus to multiply exponential numbers we simply add exponents. From this explanation you might guess, and correctly, that to divide, you subtract exponents. You can easily check this by dividing 10^5 by 10^2, which gives 1000, or 10^3. But suppose we had wanted to divide 10^2 by 10^5? Following the rule for division, we would subtract 5 from 2 and get the answer 10^{-3}, which represents the number 1/1000, or $1/10^3$. As an example, let us work out how many times larger than a hydrogen atom the sun is. In order to make this comparison, the dimensions of the two bodies *must be expressed in the same units*, so let us convert the diameter of the sun into centimeters. Since there are 100 centimeters (cm) in one meter (m), the diameter of the sun becomes $1.39 \times 10^9 \text{ m} \times 10^2 \text{ cm/m} = 1.39 \times 10^{11}$ cm.

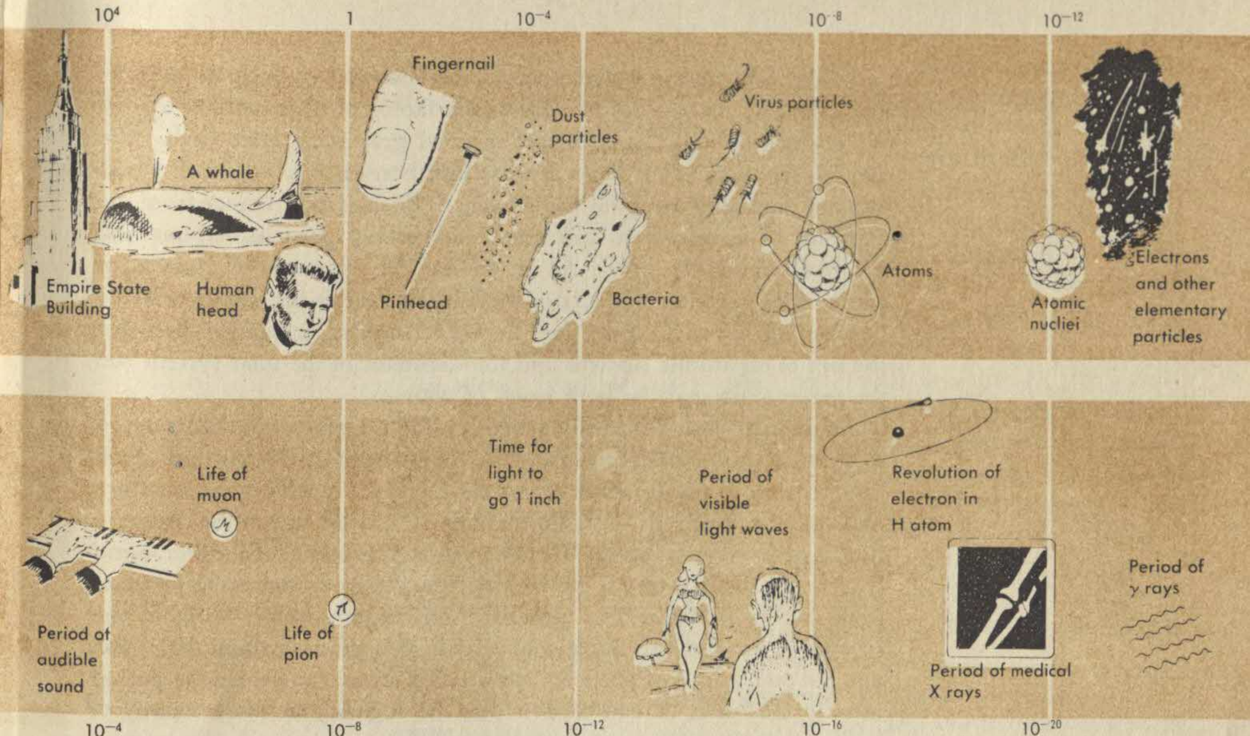


FIG. 1-1 Space and time scales of the universe.

Now we could, of course, simply divide 139,000,000,000 by 0.0000000106, but in doing this it would be very hard to keep the decimal point straight; if we use instead 1.39×10^{11} and 1.06×10^{-8}, the calculation becomes relatively simple, as we can handle the numerical multipliers and the powers of 10 separately:

$$\frac{1.39 \times 10^{11}}{1.06 \times 10^{-8}} = 1.31 \times 10^{11-(-8)} = 1.31 \times 10^{19}.$$

Sometimes, for convenience, special very large or very small units of measurement are used. Thus, in the macrocosm, astronomers often use the *light-year* as a unit to measure distances. Light travels at a speed of very nearly 3×10^8 m/sec (300,000,000 meters per second, which is about 186,000 miles/sec, or 980,000,000 miles/hr), and a light-year is the distance light travels in a year:

$$3 \times 10^8 \frac{\text{m}}{\text{sec}} \times 3.6 \times 10^3 \frac{\text{sec}}{\text{hr}} \times 24 \frac{\text{hr}}{\text{day}} \times 365 \frac{\text{days}}{\text{yr}} \times 1 \text{ yr}$$
$$= 9.46 \times 10^{15} \text{ m}.$$

Astronomers also use the *parsec*, a distance equal to 3.26 light-years. In the microcosm, we often use the *micron* (symbol μ, the Greek letter *mu*), defined as 10^{-6} m or 10^{-4} cm; and the *Ångstrom* (symbol Å), defined as 10^{-10} m or 10^{-8} cm.

In Fig. 1-1, the relative sizes of various objects in everyday life, in

the macrocosm, and in the microcosm, are shown on a decimal *logarithmic* scale, i.e., on a scale in which each factor of 10 is represented by one division. We are accustomed, in most graphs, to have each scale division represent the *addition* of some number. In Fig. 1-1, each equal distance represents an equal *multiplication*. The sizes shown range from the diameter of an electron and other elementary particles—about one hundred-thousandth of an Ångstrom—to the diameter of clusters of giant stellar galaxies. The size of the human head is just about halfway between the size of an atom and the size of the sun, or halfway between the size of an atomic nucleus and the diameter of the solar system (on the logarithmic scale in both cases, of course).

Similar vast variations will be found in the time intervals encountered in the study of the microcosm and the macrocosm. In discussing human history, we ordinarily speak about centuries; in geology, the eras are measured in hundreds of millions of years, and the age of the universe itself is believed to be about 10 billion years. The period of audible sound (that is, the time interval between two successive sound waves) ranges from 10^{-2} to 10^{-4} sec. The time required for an electron to make one revolution around the nucleus of a hydrogen atom is about 10^{-16} sec, and the oscillations of particles constituting atomic nuclei have a period of only 10^{-22} sec. A comparison (on the logarithmic scale again) of various durations encountered in the macrocosm, microcosm, and everyday life is also given in Fig. 1-1. Notice that the time interval between heartbeats is halfway between the age of our stellar system and the period of the rotation of an electron in an atom. Thus it seems that we are located about midway (logarithmically speaking) between the macrocosm and the microcosm, and can look up at the stars or down at the atoms with an equal degree of inferiority or superiority.

1-2 Units Used in the Physical Sciences

Until the beginning of the nineteenth century, the situation in the field of weights and measures was not much better than the linguistic situation at the Tower of Babel. The units of length varied from country to country, from town to town, from one profession (such as tailoring) to another (such as carpentry), and were mostly defined, rather loosely, by reference to various parts of the human body. Thus an "inch" was defined as a thumb-width, a "hand" or "palm" (still used for measuring the height of horses) as the breadth of a hand, a "foot" as the length of a British king's foot, a "cubit" as the distance from the elbow to the tip of the middle finger, a "fathom" (used in measuring ocean depths) as the distance between the tips of the middle fingers of the two hands when the arms are outstretched in a straight line, etc. In the year 1791, the French Academy of Sciences recommended the adoption of an international standard of length and suggested that the unit of length be based on the size of the earth. This unit, called the *meter*, was intended to be one ten-millionth of the distance from the pole to the equator.

The Academy prepared a "standard meter"—the distance between two fine lines engraved on a platinum-iridium bar. Later measurements of the earth, more accurate than those available to the French Academy, showed that the earth is somewhat larger than they had supposed. The exact relation between the meter and the polar circumference of the earth is of no practical importance, however, and the original length of the bar itself has been retained as the standard of length for scientific measurements everywhere. The original meter is kept at the Bureau des Poids et Mesures in Sèvres (not far from Paris), and accurate copies have been distributed to most of the countries of the world.

In recent years, it has become possible to measure lengths much more accurately by optical means than by locating two engraved lines on a metal bar. Accordingly, in 1960 the International Bureau of Weights and Measures decided to define the meter as exactly 1,650,763.73 wavelengths of the orange-red light emitted by the gas krypton-86. (Later in the book we shall see what this terminology means.) Now, if all the standard meter bars in the world were melted down into paperweights, we would still be able, by this optical definition, to reestablish the meter, and more accurately than we ever could by measuring and comparing metal bars.

Although commerce and engineering in English-speaking countries employ measurements made in feet, inches, and yards (units inherited from England along with our language), scientific measurements of length and distance are nearly always in metric units. (A peculiar situation exists at the Los Alamos Scientific Laboratory of the Atomic Energy Commission in New Mexico, where nuclear bombs are developed. The nuclear components of the bombs, which involve pure physics, are given in terms of the metric system, but the overall dimensions and the weight—which concern engineering more than science—are usually in inches and pounds.)

Along with the standard unit of length, the metric system also introduced a new unit for amount of matter, or mass. Disposing of pounds and ounces, it uses the *gram* (gm), which was intended to be equal to the mass of a cubic centimeter of water at a temperature a few degrees above the freezing point, at which it has its greatest density. (Later more accurate measurement showed the maximum density of water to be only 0.999973 gm/cm^3; but for our purposes, we can use an even 1 gm/cm^3.) A standard *kilogram* (1000 grams, abbreviated kg) was made of platinum and iridium alloy; the original is kept together with the original meter in Sèvres, and copies are distributed all over the world. While the gram and the kilogram are the standard units used in physical measurements, we also use *milligrams* (thousandths of a gram, mg) and *micrograms* (millionths of a gram, μg) to express the mass of very small amounts of matter.

Science has not yet devised a more accurate way of measuring the mass of a body, i.e., the amount of matter it contains, than by comparing

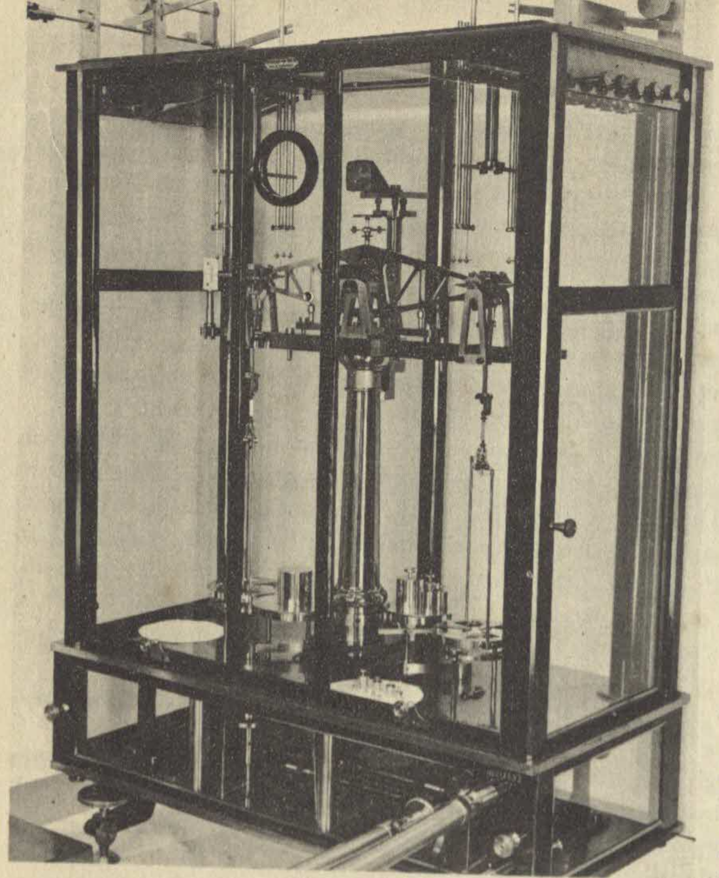


Photo courtesy National Bureau of Standards

FIG. 1–2A This balance is specially designed to compare various governmental and industrial copies with the primary United States standard kilogram, a job it can do with a precision of 7 micrograms. This gives it an accuracy greater than one part in 10^8.

it with accurate standards on a very sensitive balance. With the balance shown in Fig. 1-2A, masses can be compared with an accuracy of seven parts in 10^9. Thus all our most accurate measurements of mass are still given in terms of the actual standard kilogram.

The master clock shown in Fig. 1-2B symbolizes the third fundamental physical unit: the unit of time. A day is divided into 24 hours, each hour is subdivided into 60 minutes, and each minute further divided into 60 seconds. This system of time measurement is based upon that used in ancient Babylon and Egypt, and not even the French Revolution was able to convert it into a decimal system. In the measurement of time intervals much shorter than a second, however, the decimal system is used, and we speak about *milliseconds* (thousandths of a second, msec) and *microseconds* (millionths of a second, μsec).

Yet the definition of a second as 1/86,400 of the average length of a day throughout the year, though serving well enough to get us to breakfast on time, is not entirely satisfactory for modern scientific purposes. Astronomers have discovered that the rate at which the earth spins on its axis is not exactly constant, so that the actual length of the day varies

slightly from year to year and from century to century. In order to keep the length of the second constant for scientific purposes, our National Bureau of Standards now defines the second as the time in which it takes an atom of cesium to make 9,192,631,770 internal vibrations.

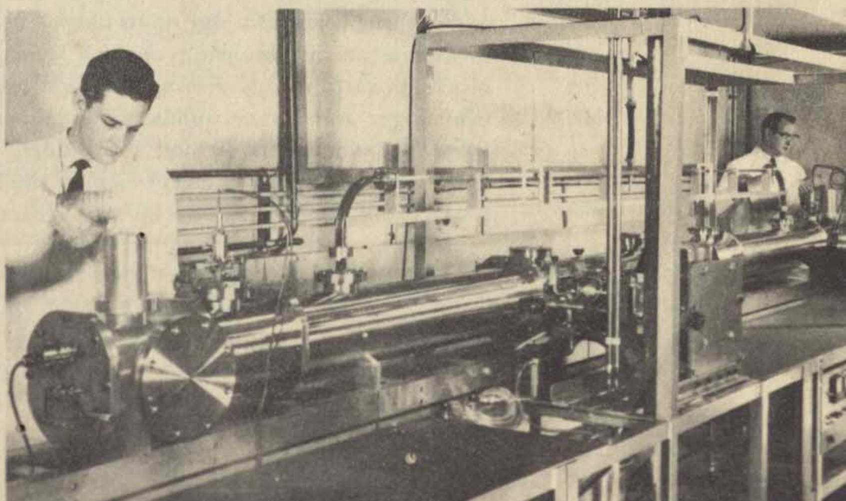
Figure 1-2B shows a cesium atomic clock that keeps time to an accuracy of 1 part in 2×10^{11}; such a clock would have to run for about 6000 years before it might gain or lose a single second.

Having defined the units for length, mass, and time, we can express through them the units for all other physical quantities. Thus one unit of velocity could be a *centimeter per second* (cm/sec), the unit of density a *gram per cubic centimeter* (gm/cm^3), etc. The above have been expressed in the system of units known as the CGS system (for centimeter-gram-second). The MKS (meter-kilogram-second) system is also coming into common use. These two decimal systems, related to each other by simple powers of 10, are accepted by scientists all over the world and represent a definite advantage over the Anglo-American system of units, in which velocity, for example, may be expressed at will in "feet per second," "miles per hour," or even in "furlongs per fortnight."

In common usage, we have a foot composed of 12 inches, a yard composed of 3 feet, and a mile that is 1760 yards or 5280 feet long (unless it is a nautical mile of 6076 feet). We use the ounce, equal to 437.5 grains (unless it is a Troy ounce, used in pharmacy and in weighing precious metals, which is 480 grains), the pound of 16 ounces (unless it is a Troy pound of 12 Troy ounces), the ton of 2000 pounds (unless it is a long ton of 2240 pounds), etc. Calculations are much easier in the

FIG. 1–2B Our measurement of time is determined by this "atomic clock" at the Boulder, Colorado, Laboratories of the National Bureau of Standards. A scientist is shown pouring liquid nitrogen (320° Fahrenheit below zero) into the device in order to help maintain the vacuum in which the cesium atoms vibrate.

Courtesy National Bureau of Standards



metric system, whose various subdivisions are all related by factors of 10 and whose relationships are indicated by the following standard prefixes, the most common of which are in boldface type:

giga	10^9	one gigaparsec	$= 10^9$ parsecs
mega	10^6	one megabuck	$= 10^6$ dollars
kilo	10^3	one kilometer	$= 10^3$ meters
hecto	10^2	one hectogram	$= 100$ grams
deca	10	one decaliter	$= 10$ liters
deci	10^{-1}	decibel	$= 10^{-1}$ $(= \frac{1}{10})$ bel
centi	10^{-2}	one centimeter	$= 10^{-2}$ meter
milli	10^{-3}	one millivolt	$= 10^{-3}$ volt
micro	10^{-6}*	one microwatt	$= 10^{-6}$ watt
nano	10^{-9}	one nanosecond	$= 10^{-9}$ second
pico	10^{-12}	one picofarad	$= 10^{-12}$ farad

The following table may be helpful in converting from one system of measurements to the other:

1 inch	= 2.540 centimeters	= 0.0254 meter
1 foot	= 30.48 centimeters	= 0.3048 meter
1 meter	= 39.37 inches	
1 mile	$= 1.609 \times 10^3$ meters	= 1.609 kilometers (km)
1 kilometer	= 0.6215 mile	
1 pound (lb)	= weight† of 453.6 grams	

weight† of 1 kilogram (kg) = 2.205 pounds

1-3 The Method Used in Physical Sciences

The observation of physical phenomena in the surrounding world leads us to the establishment of definite relations between the various quantities observed. Thus we find that the viscosity of oil depends on its temperature, the pull of an electromagnet on the strength of the current flowing through the wire, the pressure of a gas on its temperature, and the brightness of a star on its mass. Such relations, based on direct measurements, are known as *empirical laws of nature*, and the progress of observational and experimental science leads to the accumulation of ever larger and larger numbers of such empirical laws. The role of theoretical science is to find the hidden interrelations between the empirical laws and to interpret them in the light of certain hypothetical assumptions concerning the internal structure of matter, and various

* The prefix *micro* is not used with the meter, because the word *micrometer* is the name of a special instrument for measuring lengths. *Micron* is used to indicate 10^{-6} meter.

† Since the weight of a gram or a kilogram is slightly different at different locations on the earth, we must specify that this weight is to be measured at sea level, at 45° latitude.

material objects not subject to direct observation. For example, the viscosity of liquids and its dependence on temperature can be explained by a "molecular hypothesis" which assumes that all material bodies are formed by a very large number of very small particles known as molecules. The dependence of the brightness of stars on their mass can be understood if we make certain assumptions about the physical properties of the material deep in the interior of the stars and about the nature of their energy sources. In this connection, the word "model" is often used, such as "Bohr's model of the atom," or "Eddington's model of a star." It goes without saying that "model" is used here in a rather different sense than when we speak about a "model railroad" or a "model of an Indian pueblo village." A model in the physical sciences is a hypothetical description of the hidden, directly unobservable structure of certain objects, and this description is used to explain the various observed properties of such objects. Although the assumptions underlying such models and the laws that are supposed to govern them often cannot be tested by direct observation or experiment, numerous theoretical consequences can be drawn by means of mathematical reasoning and then compared with experimental and observational evidence. When the theory based on a certain model agrees with the direct empirical evidence, the belief of the correctness of the model is strengthened; and if the theory not only coincides with previously observed facts but also permits us to predict some new phenomena or regularities that are later confirmed by direct experiment or observation, we have a still stronger belief that the theory is a good one.

Questions

(1-1)

1. A man is 6 ft 1 in. tall. What is his height in meters?

2. A drill is 1/8 in. in diameter. Express this in centimeters.

3. If 325 sheets of paper make a stack 1 in. high, what is the thickness of a sheet of paper in centimeters?

4. A 31-story building is 132 m tall. What is the height of each story, in feet?

5. The diameter of a common size of European bullet is 7 mm (millimeters). Express this diameter in inches; in feet.

6. The average diameter of the earth is 7927 miles. What is this, in centimeters? in meters? in kilometers?

7. Light of a certain color is composed of a train of waves, each 5890 Å long. How many of these waves are there in an inch?

8. Light of a certain color is composed of a train of waves, each 0.437 μ long. How many of these waves are there in a foot?

9. What are the values of the following fractions?

(a) $\dfrac{6 \times 10^9 \times 4 \times 10^{-3}}{2 \times 10^{-4} \times 3 \times 10^{15}}$ (b) $\dfrac{7.19 \times 10^{-5} \times 3420}{18.0 \times 10^6 \times 2.61 \times 10^{-8}}$

10. What are the values of the following fractions?

(a) $\dfrac{8 \times 10^{-7} \times 6 \times 10^4}{3 \times 10^{-10} \times 4 \times 10^5}$ (b) $\dfrac{3.14 \times 10^8 \times 0.912}{6420 \times 1.13 \times 10^{-5}}$

(1-2)

11. The mass (which roughly means the amount of matter) of an electron is 9.11×10^{-28} gm. How many electrons would be required to make 1 gm?

12. The mass (which roughly means the amount of matter) of a proton is 1.67×10^{-27} kg. How many protons would be required to make 1 kg?

13. What is the weight of a body midway between the weight of 1 lb and the weight of 16 lb, on the ordinary, or arithmetical, scale? on the logarithmic scale?

14. What is the weight of a body midway between a weight of 1 oz and a weight of 16 lb, on the arithmetic scale? on the logarithmic scale?

15. What is the size of a creature midway between a bacterium (about 1 micron) and a small whale (about 10 meters), on the logarithmic scale?

16. What is the wavelength midway between that of visible light (about 6×10^3 Å) and a long radio wave (about 600 m), on the logarithmic scale?

17. How many μ^3 (cubic microns) are there in 1 cm^3 (cubic centimeter)?

18. How many $Å^3$ (cubic Ångstroms) are there in 1 mm^3 (cubic millimeter)?

19. What is the value of (a) $(3 \times 10^{-6})^3$? (b) of $(4.6 \times 10^2)^2$? (c) of $(46{,}000)^3$?

20. What is the value of (a) $(6 \times 10^2)^3$? (b) of $(1.2 \times 10^{-5})^3$? (c) of $(150{,}000)^2$?

21. What is the square root of: (a) 10^6? (b) 4×10^6? (c) 10×10^6? (d) 10^7? (e) 4.9×10^7? (f) 316×10^9?

22. What is the square root of: (a) 10^8? (b) 9×10^8? (c) 10×10^8? (d) 10^9? (e) 6.4×10^9? (f) 463×10^{11}?

23. What is the cube root of: (a) 10^6? (b) 8×10^6? (c) 10×10^6? (d) 10^7? (e) 2.16×10^8? (f) $\pi \times 10^8$?

24. What is the cube root of: (a) 10^9? (b) 27×10^9? (c) 100×10^9? (d) 10^{11}? (e) 6.4×10^{10}? (f) $\pi \times 10^{10}$?

chapter / two

Bodies at Rest

2-1 Equilibrium and Forces

Since we are going to study bodies in motion and what makes them move, the easiest place to start is with bodies that are not moving at all. Such objects are in a state of *static equilibrium*. As we shall see in the next chapter, any body moving in a straight line at constant speed is also in equilibrium; the stationary objects we shall consider now are merely special cases in which the constant speed happens to be zero. Let us examine a few of these special cases.

A book is probably lying on your desk—it is motionless and in static equilibrium. What do we know about the forces acting on the book? Unless there is a high wind, there are no forces tending to move it horizontally across the desk. We know, however, that there is a vertical force acting on it, trying to move it downward toward the earth's center. This gravitational attraction toward the earth is the book's *weight*; you can feel this if you hold the book in your hand. We know, too, that in order to keep the book from falling, you must push upward with a force equal to its weight. If you withdraw your hand, the downward

force of gravity acts unopposed to bring the book crashing to the floor. We can then be confident that the book lying on the table is also being acted on by two forces: the downward pull of gravity and the equal and opposite upward push of the tabletop.

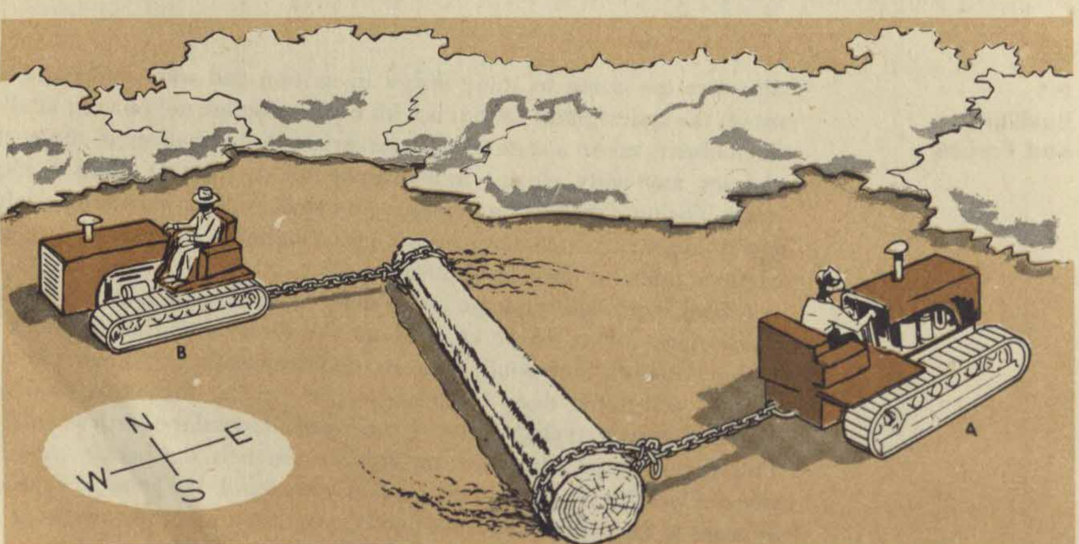
If we want to call an upward force "+" and a downward force "−", a little more elegant way of saying the same thing is to state that all the forces acting on a body in equilibrium add up to zero.

This rule of course applies to other directions as well as to up and down. If you push the book north with a force of 5 lb, the book will move unless your push is opposed by some southward 5-lb force.

2-2 Equilibrium and Torques

Even if we were sure that all the forces acting on a body added up to zero, this alone would not guarantee that the body would be in equilibrium. In Fig. 2-1, the log on the ground is being acted on by two equal and opposite forces. Tractor A pulls one end of the log east with a force of 2000 lb; tractor B pulls the other end west with an equal 2000-lb force. These two forces add up to zero, but we know the log will *not* lie motionless; instead, it will begin to rotate, as shown. The rotating effect of a force depends, not only on how big the force is, but also on where it is applied. In Fig. 2-2A, for example, it will be very hard to turn the rusted nut. The turning effect, or *torque*, is the force F multiplied by the distance d_1. By slipping a piece of pipe over the wrench handle, as in Fig. 2-2B, the distance, or *lever arm*, is increased to d_2, and

FIG. 2–1 Two equal and opposite forces exert a torque on a log.



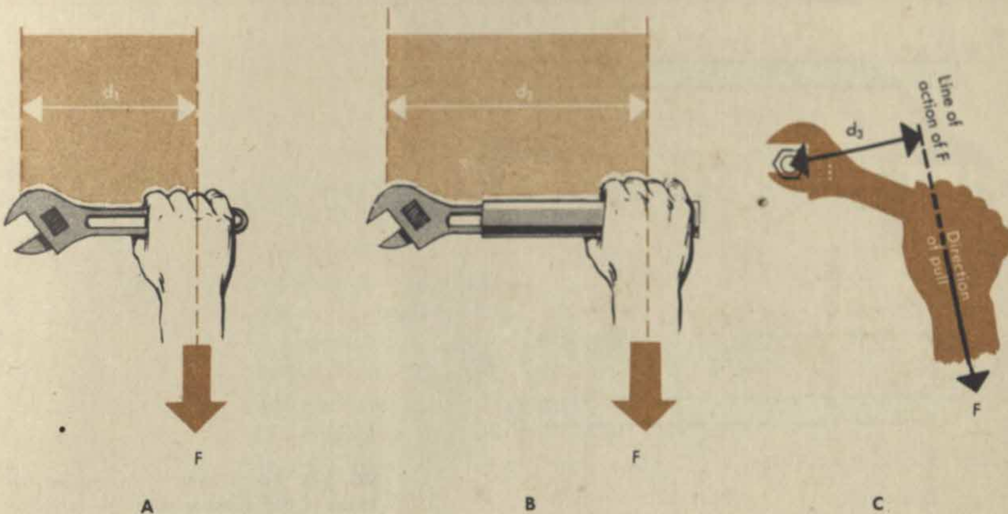


FIG. 2–2 The torque produced by a force depends on the force and on its lever arm.

the torque Fd_2 is made much greater without any increase in the force.

We must note here that the lever arm is always measured from the center of rotation *perpendicular to the line of action of the force F*. This is seen to be the case for d_1 and d_2 in Fig. 2-2A and B, in which the pull *is* perpendicular to the handle of the wrench. In Fig. 2-2C, however, the pull is at an angle, and the situation is different. The distance d_3 (which is the lever arm) must be measured from the center of the nut in the direction shown, perpendicular to the dashed *line of action* of the force F, i.e., the prolongation of the arrow or vector representing the pull of the hand.

Figure 2-3 shows a light bar with weights suspended from its ends, balanced on a narrow support at B. If we were to balance such a bar with weights as shown, we would find that the distances BC and AB would have to be in the same proportion as the weights hung from A and C, respectively; in this particular case, in the proportion, or ratio, of 3 to 1. If we imagine the bar trying to turn about B, force A will have a torque of $3 \times 1 = 3$ ft-lb counterclockwise, and will be just balanced by the clockwise torque of force C, which equals $1 \times 3 = 3$ ft-lb. (Since we obtain torque by multiplying a distance times a force, the units, or dimensions, of torque are always expressed as distance-force units—in this case, feet $\times$ pounds, foot-pounds.) The force at B will have no torque about B as a center of rotation, because its lever arm is zero.

In dealing with the torques in this example, we did not need to know the upward force at B. It would, however, be very easy for us to find, since we know that in equilibrium the forces up must equal the forces down: $F_B = 3 + 1 = 4$. If, using this knowledge of the force at B, you check the torques around A or C or any other point, you will get the

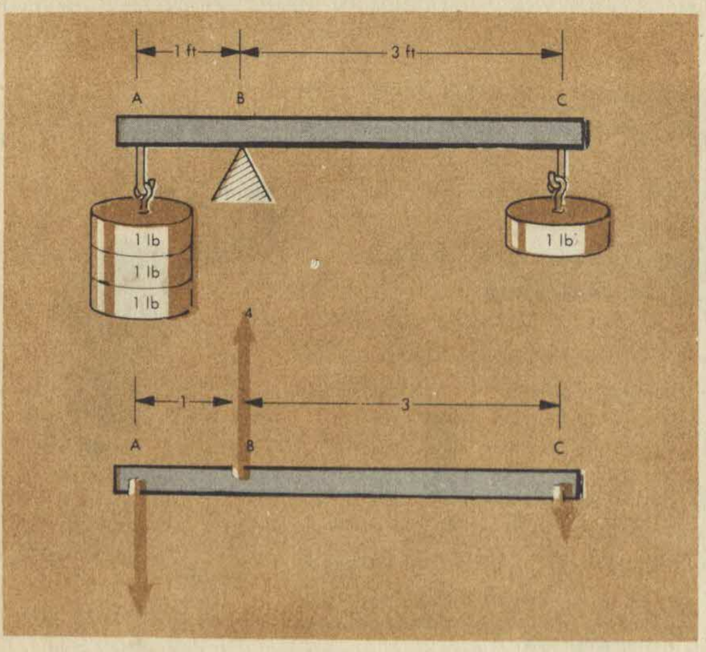


FIG. 2-3 For a body in equilibrium, the forces in one direction equal the forces in the opposite direction, and the clockwise torques equal the counterclockwise torques.

same results—namely, that the clockwise torque equals the counterclockwise.

Using + and − signs to indicate upward and downward forces, and clockwise and counterclockwise torques, we can put the conditions for equilibrium into the following very brief and simple form:

$$\sum F = 0 \quad \text{and} \quad \sum T = 0.$$

The $\sum$ is the Greek capital letter *sigma* and is used by mathematicians to mean "the sum of." These simple formulas can be checked out as follows for the lever shown in Fig. 2-3:

$$\sum F = 4 - 3 - 1 = 0$$

$$\sum T = 3 \times 1 - 1 \times 3 = 0 \quad (\text{around } B)$$

or

$$\sum T = -4 \times 1 + 1 \times 4 = 0 \quad (\text{around } A)$$

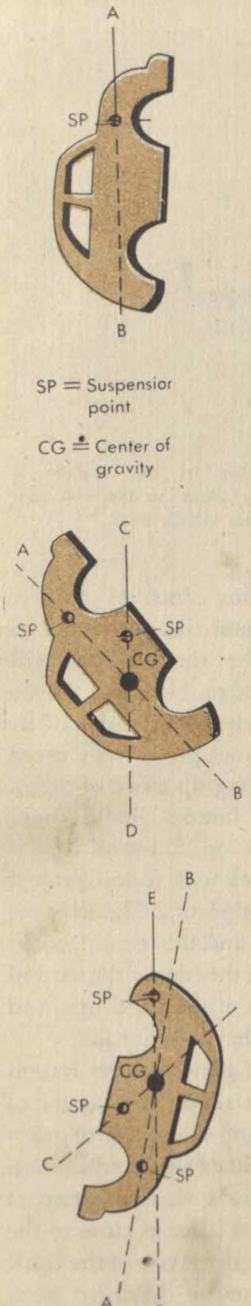
or

$$\sum T = -3 \times 4 + 4 \times 3 = 0 \quad (\text{around } C).$$

2-3 Center of Gravity

In solid objects that possess a definite shape, we can single out an important point known as the *center of gravity* (C.G.). The force of gravity acts, of course, on all parts of any given object, but the center of gravity can be defined by saying that objects subjected to gravity behave as if there were only a single force applied at that point. If a body is supported at its center of gravity, it will be in balance and have no tendency to move or rotate in any direction.

FIG. 2-4 A simple experimental method for locating center of gravity of an irregular figure.



For any symmetrical body of uniform composition, the center of gravity is the same as its geometrical center. In the case of an object of more complicated shape, the center of gravity can always be found by suspending it on a string attached first at one point and then at another point on its surface. An object always comes to rest with its center of gravity directly under the point of suspension. Suppose we cut an object out of plywood with the shape shown in Fig. 2-4. If we suspend it at point A, it will hang in the way shown in Fig. 2-4 (top), and its center of gravity must be located somewhere on the line AB. If we suspend it at another point C, the object will hang as shown in Fig. 2-4 (center), and the center of gravity must be somewhere on the line CD. Thus the exact location of the center of gravity is determined by the intersection of the lines AB and CD. It is unnecessary to suspend it from a third point E (bottom), because the vertical line EF must also pass through the intersection of the previous two lines.

In our investigation in the preceding section, concerning the light bar with the weights at its ends, the center of gravity of the whole arrangement was at the balance point B. In this problem, we took advantage of the word "light" to ignore the effect of gravity on the bar itself. Most real bars and sticks, however, are heavy enough so that their own weights must also be taken into account if we are to expect accurate answers. Let us go back to this example, as illustrated in Fig. 2-3, and assume that the bar itself weighs 1 lb. Now, where will the balance point be, with the weights still attached as they were?

Figure 2-5 shows the same bar, with the addition of the force marked W, representing the weight of the bar itself. Although gravity pulls equally along the entire length of the bar, we can consider the entire pull of 1 lb to be exerted at the center of gravity of the bar. We can assume that the bar is uniform, so that its center of gravity will be at its center, 2 ft from each end. Although the upward supporting force at B can be easily calculated to be 5 lb (from $F_{up} = F_{down}$), we do not know where it will be located. Assume it to be x ft from, say, the left end. Then, with the bar balanced and in equilibrium, the total torques of all the forces acting on the bar will add up to zero, no matter what point we choose to pick as the center of rotation. Point A, at the left end, is as good as any. So, using A as our center and putting torques clockwise = torques counterclockwise, we get

$$3 \times 0 + 1 \times 2 + 1 \times 4 = 5 \times x$$

$$5x = 6$$

$$x = \tfrac{6}{5} = 1.2 \text{ ft.}$$

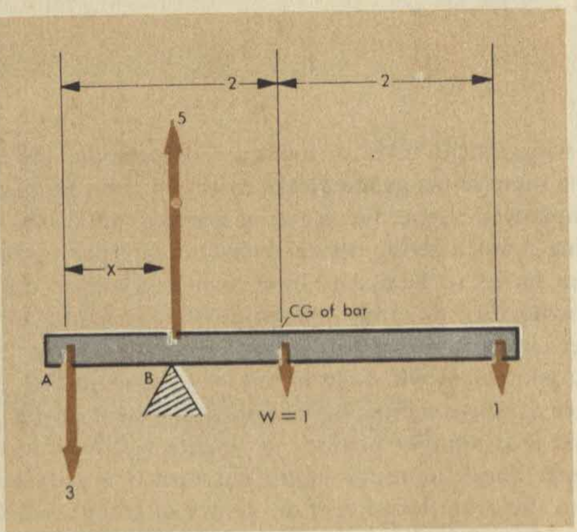


FIG. 2-5 Including its own weight in balancing a bar similar to the one considered weightless in FIG. 2-3.

Let us consider the following problem concerning center of gravity. Suppose we have a large number of books of equal size and want to pile them at the edge of a table in such a way that the top book will protrude as much as possible beyond the table's edge. How can we do it? If the bottom book is placed so that it protrudes by almost half its length, the second and all the following books cannot project any more without falling, and nothing will be gained by piling up more of them. Instead of starting the arrangement with the bottom book, then, suppose that we first consider the book that is on top. Since all that is required from it is that it not fall from the pile below it, it can project out by just a little less than half its length. By inspecting Fig. 2-6, we see that the common center of gravity of the first and the second books will be located a quarter-book length to the right of the edge of the second book. Thus, if these two books are placed on top of the third one and overhang it a quarter-length (or just a bit less), they will not fall.

Let us go one step further and find the center of gravity of the system of the three top books. The way the books are stacked, the center of gravity of the first two is halfway between them, and the center of gravity of the third book is, of course, in its middle. Since the combination of the first two books is twice as heavy as the third book, we should expect the center of gravity of all three books to be located twice as close to the center of gravity of the first two as it is to the center of gravity of the third one. A glance at Fig. 2-6 shows that the overhang of the third book should be one-sixth of its total length. If we proceed in the same way down the pile, we find that the next two overhangs will be one-eighth and one-tenth, respectively. With five books, the distance of the outer edge of the top book from the edge of the table will be

$$\left(\tfrac{1}{2} + \tfrac{1}{4} + \tfrac{1}{6} + \tfrac{1}{8} + \tfrac{1}{10}\right) \text{ book lengths} = 1.14 \text{ book lengths.}$$

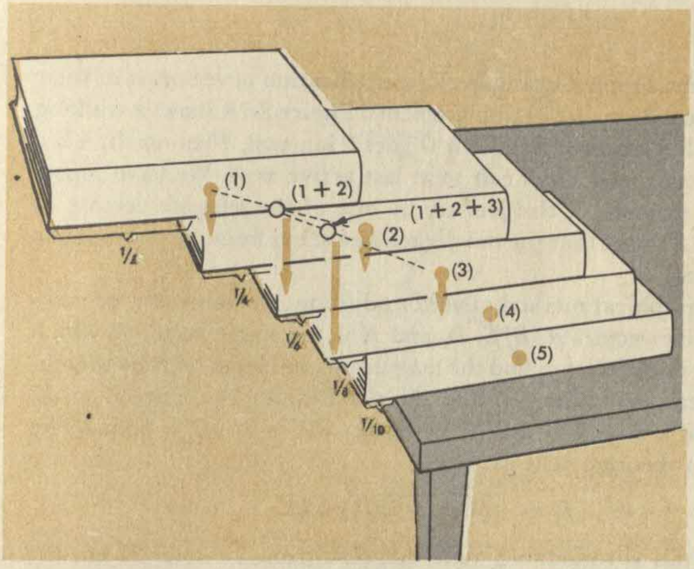


FIG. 2–6 The best way to pile books to extend them beyond the edge of a table. The point (1 + 2) is the center of gravity of the top two books. The point (1 + 2 + 3) is the center of gravity of the three top books, and is located one-third of the way between the double weight applied at (1 + 2) and the single weight applied at (3).

Thus, by piling books up in a rational way, we can do much better than a half-book overhang—in fact, better than a full-book overhang. If we use more than five books, the sum in the bracket above must be extended by adding $\frac{1}{12}$, $\frac{1}{14}$, $\frac{1}{16}$, etc., and it can be proved mathematically that the sum of a series of terms of this kind can become as large as we want, provided that we add enough terms. By stacking an unlimited number of books, therefore, we can make the top book protrude any desired distance beyond the edge of the table. Because of the rapidly decreasing contribution of each new book, however, we would need the entire Library of Congress to make the overhang equal to three or four book lengths!

2-4 Vectors

Some physical quantities, such as age, amounts of money, and temperature, have magnitude only—that is, they can be described by a single number that tells the amount of the property we are interested in: 30 years, \$27.19, 75°, etc. Such quantities are called *scalars*. However, there are other quantities that have direction as well as magnitude. We cannot describe a force, for example, merely by saying that it is equal to 30 lb. The direction of the force must also be given if we are to have a complete picture of it. Such quantities are called *vectors*. Vector quantities can be very conveniently represented by arrows. The length of the arrow, drawn to some convenient scale, represents the magnitude of the quantity, and the arrow is pointed in a direction parallel to the direction of the vector quantity.

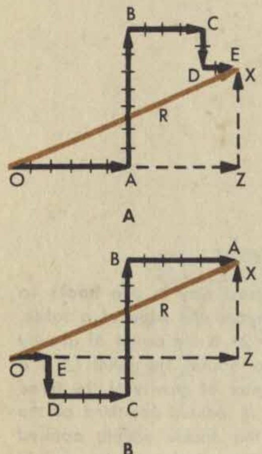
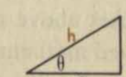
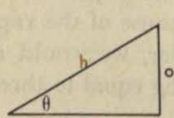


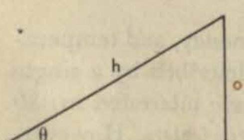
FIG. 2-7 The graphical addition of vectors.



$$\sin \theta = \frac{\text{side opposite } \theta}{\text{hypotenuse}} = \frac{o}{h}$$



$$\cos \theta = \frac{\text{side adjacent } \theta}{\text{hypotenuse}} = \frac{a}{h}$$



$$\tan \theta = \frac{\text{side opposite } \theta}{\text{side adjacent } \theta} = \frac{o}{a}$$

FIG. 2-8 Trigonometric functions of an angle.

Probably the simplest example of the application of vectors is in their use to represent distances or displacements. Figure 2-7A shows a walking tour in which a man starting from O goes 6 km east, 7 km north, 4 km east, 2 km south, and 2 km east to at last arrive at X. We have represented each segment of this journey by one of the separate vectors A, B, C, D, and E; note that the tail of each one starts from the point of the preceding one.

This is a graphical method of vector addition; the *vector sum*, or *resultant*, of the five vectors A, B, C, D, and E is the single vector R which goes from O to X. We can find the magnitude, or length, of R by sketching in the east-west line OZ and the north-south line from Z to X. The line OZ is $6 + 4 + 2 = 12$ km long; ZX is $7 - 2 = 5$ km. The Pythagorean theorem tells us that

$$R = \sqrt{(12)^2 + (5)^2} = 13.$$

(Note that this is something quite different from the total 21 km the man has walked.)

It is a good thing to know that it makes no difference in what order vectors are added. Figure 2-7B shows them added in the order E, D, C, B, A. A quick check will show that OZ (the east-west *component* of R) is again 12, that ZX (the north-south *component* of R) is 5, and that R is identical with the R of Fig. 2-7A.

We now have half the information we need to describe a vector. It would be convenient to be able to say also that "R points in a direction so many degrees north of east." Of course, we could lay a protractor on the figure (if it is drawn accurately to scale) and thereby find the angle ZOX to be about 23°. But it is generally desirable to use some more accurate and more convenient analytical method for finding out about angles. Fortunately for us, the patient labors of the men who computed the tables of trigonometric functions have made this easy, and a condensed table is given on page 573.

Figure 2-8 shows the right-triangle relationships that define the three principal trigonometric functions—the *sine* (abbreviation: sin), the *cosine* (cos), and the *tangent* (tan). The angle θ (the Greek letter theta) is the same in all the various-sized triangles shown, which means that all the triangles have the same shape. Therefore, the ratio o/h or a/h or o/a will be the same, regardless of the size of the triangle; in other words, these ratios depend *only* on the angle θ, which determines the shape of the triangle.

We can now return to Fig. 2-7 and see that the tangent of angle $ZOX = \frac{5}{12} = 0.416$; $\sin ZOX = \frac{5}{13} = 0.384$, and $\cos ZOX = \frac{12}{13} = 0.923$. From any one of these, we can find from the table (or from a slide rule) that angle $ZOX = 22.6°$, and we can thus give the direction of R as 22.6° north of east.

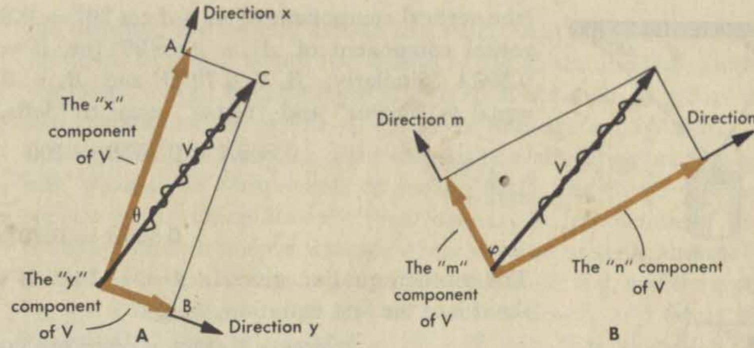


FIG. 2-9 Resolution of a vector into components.

For many purposes, it is useful to be able to replace a vector by two or more *components*—that is, two or more vectors which will add together to give the original vector. Almost always we shall want these components to be at right angles to one another, and Fig. 2-9 shows how this can be done in a process called *resolving* a vector into components. We see that the x component of V is $V \cos \theta$ and that the y component (which is equal to AC) is $V \sin \theta$. Similarly, if we had for some reason chosen different directions, we would have V's m component equal to $V \cos \phi$ and its n component equal to $V \sin \phi$. The original vector V is shown marked out because it has been replaced by its components, which add up vectorially to produce the same effect as V itself.

2-5 Oblique Forces

The use of vectors and their components makes it possible for us to put the idea of equilibrium on a more secure and definite footing. In the simple examples we have used so far, the applied forces have always been parallel, so that it was easy to see that the "ups" just nullified the "downs" and that the "lefts" exactly counterbalanced the "rights." For torques, also, we heretofore have considered only forces which were perpendicular to given lever arms, so that the torques were easy to calculate. Most actual examples, though, include forces and torques that are not so considerately arranged.

For example, look at a weight suspended from ropes as shown in Fig. 2-10A. The pulls of the ropes on the knot P will be along the directions of their lengths, as indicated by vectors A and B in Fig. 2-10B. Also acting on the knot is the pull of gravity—100 lb straight down. By resolving A and B into vertical and horizontal components, we can as simply as before deal with "ups" and "downs" and "lefts" and "rights." A_v

(the vertical component of A) is $A \cos 30° = 0.866A$, and A_h (the horizontal component of A) is $A \cos 60°$ (or, if you prefer, $A \sin 30°$) $= 0.500A$. Similarly, $B_v = 0.707B$ and $B_h = 0.707B$. Putting "ups" equal to "downs" and "rights" equal to "lefts," we have

$$0.866A + 0.707B = 100$$

and

$$0.500A = 0.707B.$$

The second equation gives us $A = 1.414B$. If we substitute this value for A into the first equation, we get

$$0.866 \times 1.414B + 0.707B = 100$$
$$1.932B = 100$$
$$B = 51.76 \text{ lb.}$$

Substituting this value for B back into one of our first equations, we find that

$$A = 73.19 \text{ lb.}$$

The example we have just considered was one in which all the forces acted together at a single point, so that it was not necessary to deal with torques. In many cases, however, it will be necessary to compute the torque of an oblique force. Figure 2-11A shows a hinged wall shelf used as a desk. The shelf itself weighs 20 lb, and a 40-lb typewriter is placed on the shelf with its center of gravity 10 in. from the front edge. The shelf is supported on each side by a chain fastened as indicated in

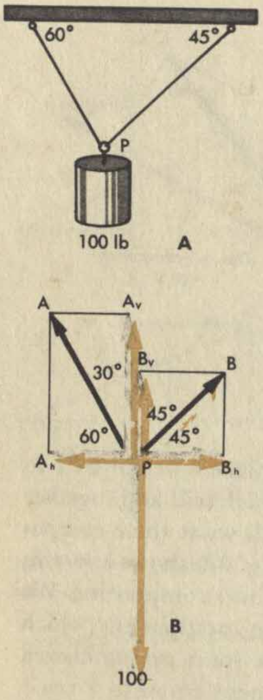
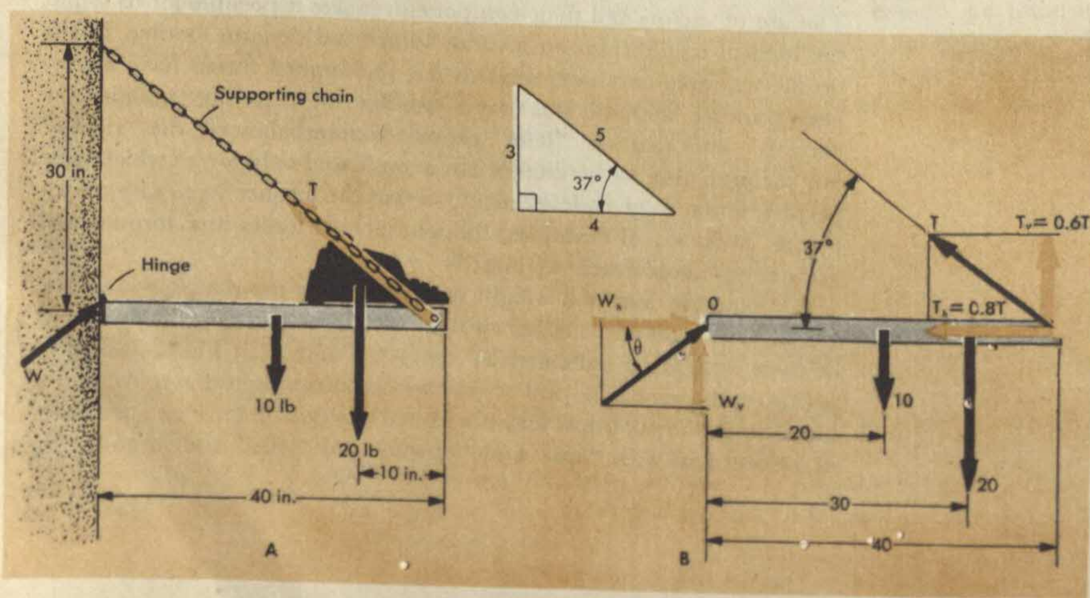


FIG. 2-10 Resolution of oblique vectors into vertical and horizontal components.

FIG. 2-11 Resolution of a force into components at right angles to one another.



the drawing, and we can assume that each chain and hinge carries half the load. In order to find the tension in the chain and the reaction at the hinge, we can start by *isolating* one end of the shelf, as in Fig. 2-11B, and marking on the sketch all the forces that are acting on it.

Forces W and T can be dealt with quite easily if we break them up into rectangular components as shown. From the given dimensions, we can see that the pull of the chain is along the hypotenuse of a 3–4–5 triangle, which is a right triangle whose smallest angle is about 37°. As can be seen from the drawing, the sine of this angle is $\frac{3}{5}$, and the cosine $\frac{4}{5}$. Figure 2-11B shows us (forces up = forces down, and forces left = forces right) that $W_v + 0.6T = 10 + 20$ and that $W_h = 0.8T$, but this is not enough information to let us solve for any of the three unknowns. For further information we must turn to torques. Because the shelf is stable and stationary, it is not turning about any axis whatever; this means that around any axis we choose, the sum of the clockwise torques must equal the sum of those acting to turn it in a counterclockwise direction. Since we are thus free to choose any axis we like, let us exercise this choice intelligently to save as much work as we can. A look at Fig. 2-11B tells us that it will be to our advantage to choose point O, the hinge, since three of our four force components pass through O, and thus have zero torque about it. Therefore, setting the torques tending to turn the shelf clockwise about O equal to those acting in a counterclockwise direction, we have

$$10 \times 20 + 20 \times 30 = T_v \times 40$$

so that

$$T_v = 20 \text{ lb}$$

and

$$T = \frac{20}{0.6} = 33.3 \text{ lb.}$$

Substituting this value back into the force equations, we get

$$W_v = 10 \text{ lb}$$

and

$$W_h = 26.7 \text{ lb.}$$

If we also want the direction of the hinge reaction, we see that $\tan\theta = W_v/W_h = 0.375$, and so $\theta = 20.55°$.

2-6 Friction

So far, we have ignored one of the most important classes of forces in the world—the forces of friction. We all realize that if it were not for the friction between our shoes and the floor, we could not walk; cars move and stop only through the forces of friction between tires and road.

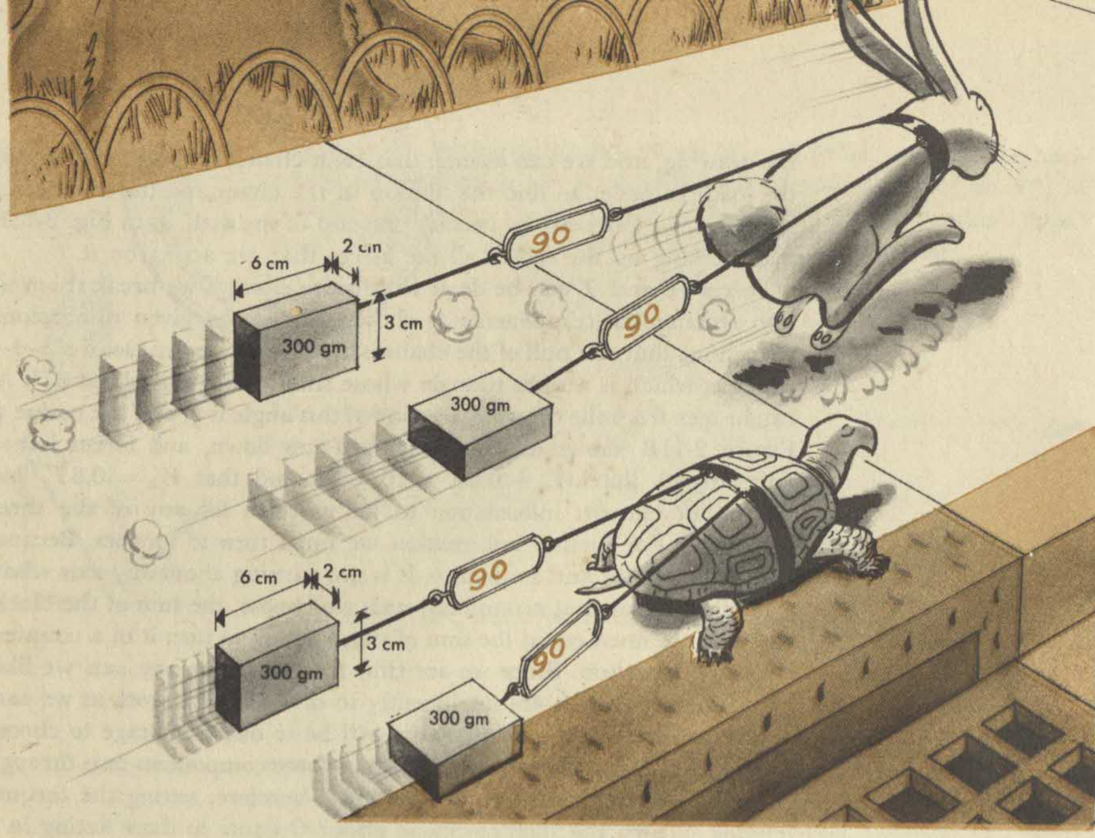


FIG. 2-12 The position of the block has no effect on the force of friction, which depends only on the nature of the surfaces and the force pressing them together (here, the weight of the block).

(We are reminded of this in winter when roads are icy and these frictional forces very small.) Experiment shows that the frictional force between any two sliding surfaces is directly proportional to the force with which the surfaces are pressed together. The proportionality constant is called the *coefficient of friction* between the surfaces, and in most textbooks is assigned the symbol μ (the Greek letter *mu*), so that

$$F_{fr} = \mu \times F$$

The value of μ depends on the materials that are sliding and is only very slightly affected by other factors, such as speed or contact area.

Let us take an iron block 2 cm by 3 cm by 6 cm and slide it along a concrete walk. The weight of the iron block is about 300 gm, and the coefficient of friction between iron and concrete is about 0.3. From this we can compute that the force needed to drag the iron block will be $300 \times 0.3 = 90$ gm. We shall find that it makes almost no difference if the iron block stands on end or lies flat on the walk; and whether we move it along rather slowly or at a higher speed, it will still take a 90-gm pull (see Fig. 2-12).

A simple experiment that demonstrates the action of friction can be carried out with a ruler or a golf club. Take a ruler and support each end by your index fingers. Now move your hands closer and closer together and note that the stick will slide over your right finger, then over the left one, then over the right, then over the left again, etc. Finally, when your two fingers come together, the ruler will still be in equilibrium, and its middle point (which in this case, of course, is the center of gravity) will be located between the two fingers. If you repeat the same experiment with a golf club, which has a heavy piece of iron fastened to one end, you will find that the center of gravity is located closer to the heavy end of the club. The alternating sliding of the ruler or the golf club over your fingers is caused by the fact that the friction force between any two objects sliding on one another is larger the more strongly they are pressed together, and by the fact that the finger located closer to the center of gravity is supporting the larger fraction of the total weight of the ruler or the club. The object will always slide on the finger supporting the least weight, which is the one farthest from the center of gravity. As the golf club slides first on one finger and then on the other, its center of gravity will always remain between them.

As another example, consider a man climbing a 40-lb, 13-ft ladder with its base 5 ft out from a smooth vertical wall, as shown in Fig. 2-13A. What must be the coefficient of friction between the foot of the ladder and the ground, if the man is to be able to climb to the top without the

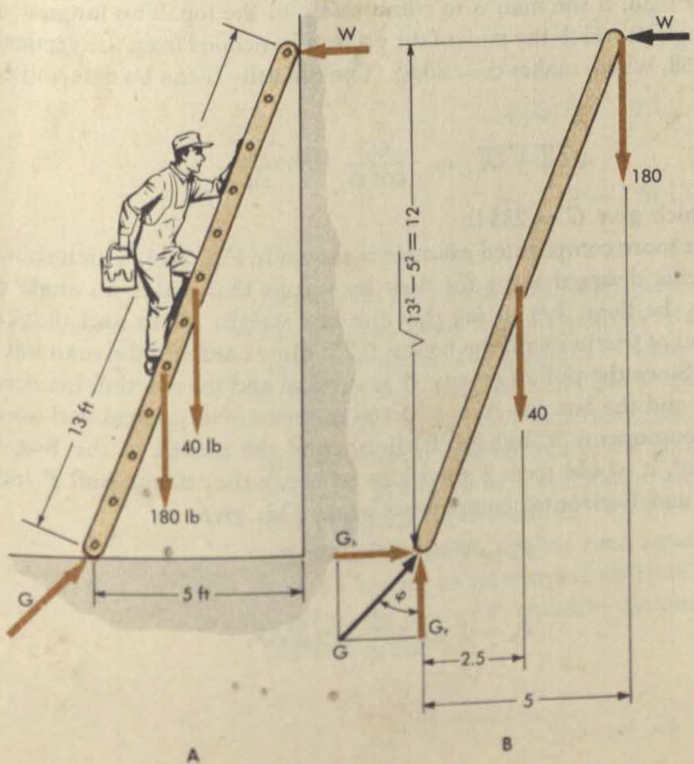


FIG. 2-13 Forces and torques acting on an inclined ladder.

ladder slipping? Figure 2-13B isolates the ladder and all the forces acting on it. Because the wall is smooth, no frictional force can be transmitted between wall and ladder, and W is thus perpendicular to the wall. The 180-lb weight of the man is shown at the top, and if we assume the ladder to be uniform, its 40-lb weight acts at its center. The push of the ground G is unknown in both direction and magnitude, and it will be convenient to break it into components G_v and G_h. From Fig. 2-13B, we can immediately write

$$G_v = 40 + 180 = 220$$

and

$$G_h = W.$$

The best choice for the axis of torques is the base of the ladder, because around this axis both G_v and G_h have zero torque and are thus eliminated:

$$40 \times 2.5 + 180 \times 5 = W \times 12$$

or

$$W = 83.3 \text{ lb}.$$

This value, substituted back into the force equations, gives us

$$G_h = 83.3 \text{ lb}.$$

The coefficient of friction at the base of the ladder cannot be less than G_h/G_v, or 0.38, if the man is to climb safely to the top. The tangent of ϕ, the angle at which the push of the ground is inclined from the vertical, is also 0.38, which makes $\phi = 20.8°$. The quantity G can be determined as

$$\sqrt{G_v^2 + G_h^2} \quad \text{or} \quad \frac{G_v}{\cos\phi} \quad \text{or} \quad \frac{G_h}{\sin\phi}$$

all of which give $G = 235$ lb.

A little more complicated example is shown in Fig. 2-14, which shows a box being dragged along the floor by a rope that makes an angle of 30° with the floor. Let us say that the box weighs 500 lb and that its coefficient of friction with the floor is 0.25. How hard will the man have to pull? Since the pull of gravity W is vertical and the reaction between the floor and the box can be considered in terms of its vertical and horizontal components R and F_{fr} (friction), and the motion of the box is horizontal, it would seem a good idea to break the oblique pull P into vertical and horizontal components also. This gives us

$$P_v = P \sin 30° = 0.5P$$

and

$$P_h = P \cos 30° = 0.866P.$$

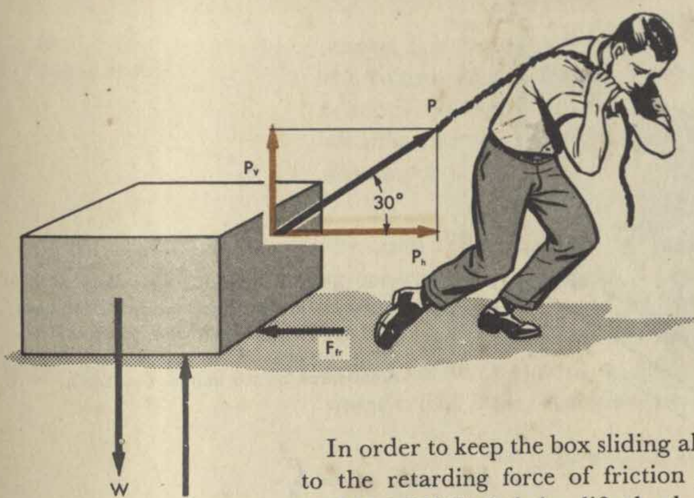


FIG. 2-14 The oblique force P both pulls the box along and reduces the force with which the box presses against the floor.

In order to keep the box sliding along the floor, P_h need be only equal to the retarding force of friction F_{fr}. This frictional force is $0.25R$ (NOT $0.25W$! P_v helps lift the box upward, so that $R = W - P_v$). We can now write quite simply

$$\begin{aligned} P_h &= F_{fr} \\ &= 0.25R \\ &= 0.25(W - P_v). \end{aligned}$$

From this, we get

$$\begin{aligned} 0.866P &= 0.25(500 - 0.5P) \\ &= 125 - 0.125P \end{aligned}$$

or

$$0.991P = 125$$

and

$$P = 126.1 \text{ lb.}$$

2-7 Fluid Pressure

Liquids do not have a definite shape of their own and readily assume the shape of any vessel they are poured into. But in doing so, a liquid retains its total volume, and a gallon of water will remain a gallon whether it is poured into a flat dish or into a tall, narrow container.

Consider water in a cylindrical glass. Because of its weight, the water exerts a force on the bottom of the glass, a force the same as that which would be produced by a cylindrical piece of ice if the water were frozen and the glass walls removed. This force is distributed over the entire bottom of the glass, so that each square centimeter of the area of the bottom carries its own equal share of the load. This *force per unit area* is called the *pressure* P, equal to the total force divided by the area over which it is exerted:

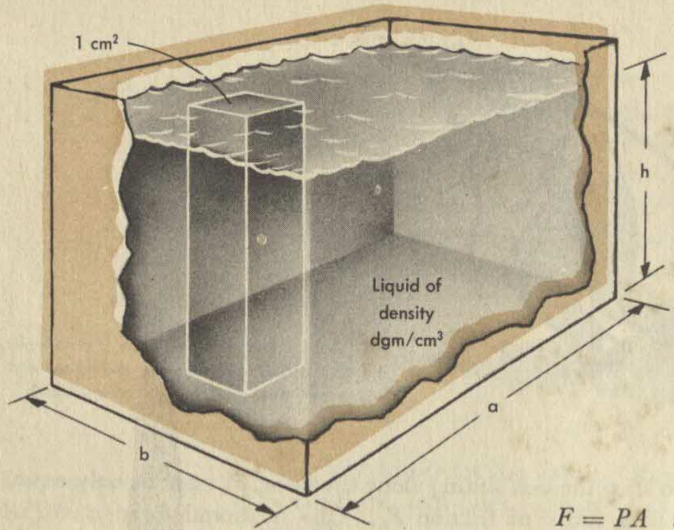


FIG. 2-15 Each square centimeter of the bottom of the tank supports a column of liquid that weighs hd grams. This pressure (hd) times the area of the bottom (ab) is the total force on the bottom ($=abhd$).

$$F = PA \quad \text{and} \quad P = \frac{F}{A}.$$

All equations must be equations in every sense of the word—not only must the numerical values be the same on both sides, but so must the dimensions, or units in which the various quantities are measured. If, for example, we have a pressure of 10 lb/in² exerted on a surface whose area is 3 ft², we would get the following for F:

$$F = PA = \frac{10 \text{ lb}}{\text{in}^2} \times 3 \text{ ft}^2 = \frac{30 \text{ lb-ft}^2}{\text{in}^2} = ???$$

These units tell us nothing at all, and we see that in order to get a sensible answer we must either convert the in² into ft² or vice versa. If we do this, we can then cancel the area units in the numerator with the similar ones in the denominator and come out with a reasonable answer of so many pounds. Let us choose to convert the area into square inches: since there are $12^2 = 144$ in² per ft², the area is $3 \times 144 = 432$ in², and we have

$$F = \frac{10 \text{ lb}}{\text{in}^2} \times 432 \text{ in}^2 = 4320 \text{ lb}.$$

Or, if we prefer, we can convert the pressure: since $P = 10$ lb/in² means that there is a force of 10 pounds on every square inch, the force on each square foot will be 144 times as great, and the pressure can be expressed as $10 \times 144 = 1440$ lb/ft². Then

$$F = \frac{1440 \text{ lb}}{\text{ft}^2} \times 3 \text{ ft}^2 = 4320 \text{ lb} \quad \text{as before.}$$

Figure 2-15 shows a tank containing an imaginary column of liquid 1 cm square and extending h cm from the bottom to the top free surface. The volume of this column is h cm³, and its weight will be h times the weight of a single cubic centimeter. This latter figure, the weight of 1 cm³ (or 1 ft³ or 1 m³) of any substance, is called the *density* of the sub-

stance, and we can let it be represented by the letter d. The weight of the column of liquid then will be the weight of hd gm, and this weight is supported by 1 cm^2 of the bottom. Since force per unit area is what we mean by *pressure*, we can use this general formula for the pressure at any depth in a liquid:

$$P = hd.$$

We used the bottom of the tank in making our calculation, but the same arguments will hold for the pressure at any depth. Here too, as in any quantitative work, we must keep our eyes on the units. As an example, we can calculate the pressure due to the overlying water at the bottom of the ocean at a depth of exactly 2 mi. The density of sea water is 64.0 lb/ft^3, and if we merely substitute in the equation, we have

$$P = hd = 2 \text{ mi} \times \frac{64 \text{ lb}}{\text{ft}^3} = \frac{128 \text{ lb-mi}}{\text{ft}^3}.$$

which makes as little sense as force expressed in $\text{lb-ft}^2/\text{in.}^2$. But if we convert the 2 mi into $2 \times 5280 = 10{,}560$ ft, the answer will be perfectly reasonable:

$$P = hd = 10{,}560 \text{ ft} \times \frac{64.0 \text{ lb}}{\text{ft}^3} = 676{,}000 \text{ lb/ft}^2.$$

Water, or any other fluid, since it does not have a rigid shape, cannot resist a pressure exerted on it in only one direction, the way, for example, a block of steel can when it resists being squeezed in a vise. Instead, liquids squash out in all directions and so exert equal pressure in all directions. Thus our formula $P = hd$ is equally useful in calculating the pressure against the wall of a container at any depth, no matter at what angle the wall happens to be. For this reason, the shape of the container makes no difference. In Fig. 2-16, if both vessels are filled with a liquid of density d, the pressure at the bottom is h_1d for both, and at the points marked A the pressure in both is h_2d.

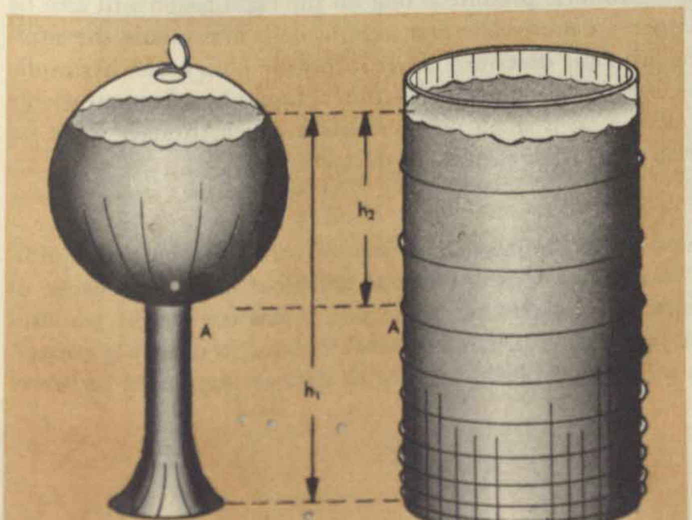


FIG. 2-16 The pressure in vessels of different shape is the same at equal depths of the fluid in them.

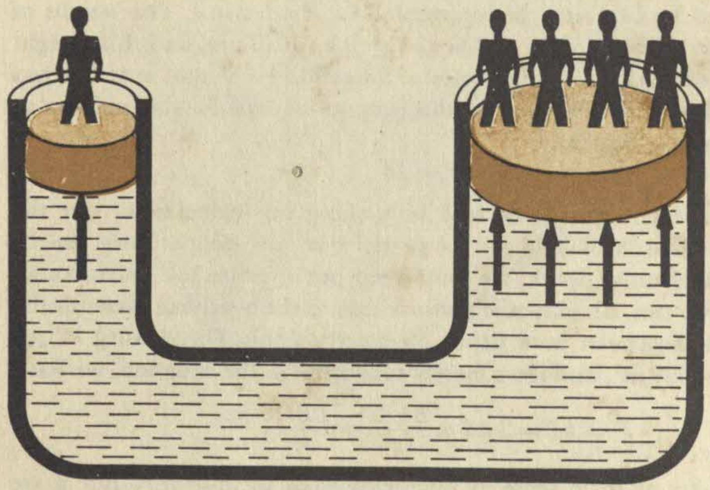


FIG. 2-17 A demonstration of Pascal's Law.

2-8 Pascal's Law

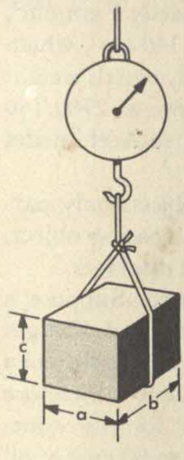
Since the pressure on any small piece of a fluid is the same in all directions, ***any increase of pressure on a fluid in a closed container will be transmitted equally to every part of the fluid.*** This basic law was discovered by the French physicist Blaise Pascal (1623–1662) and carries his name. Imagine a closed vessel with two vertical cylinders of different diameters protruding from its upper part (Fig. 2-17). These cylinders are fitted with pistons that can be loaded with various numbers of heavy weights. If we place one weight on the piston in the narrower cylinder, it will produce a pressure within the liquid, and this same pressure will be transmitted to all parts of the vessel, including the surface of the larger piston. Since, however, the area of that piston is larger, the total force acting on it will be larger, too. In the example shown in Fig. 2-17, the cylinder on the right is twice as large in diameter, so the areas of the two pistons stand in the ratio of 4 to 1. Therefore, since the total *force* of hydrostatic pressure acting on the right piston will also be four times larger, we must place four weights on it to maintain the equilibrium. The principle described above forms the basis of the hydraulic press in which the pressure created within a liquid by a comparatively small force acting on a small piston exerts a much stronger force on another piston of considerably larger diameter.

2-9 Archimedes' Law

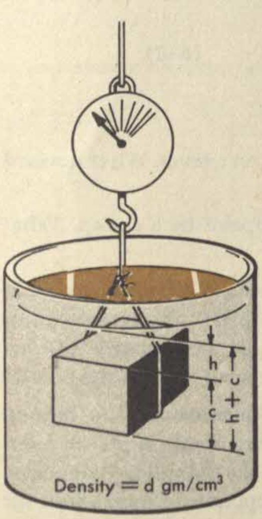
We turn now to the important subject of solids floating in liquids. Everybody knows that a piece of wood will float in water because its density is smaller than that of water (i.e., it has less weight per unit volume) and that a piece of metal will sink because its density is greater. Although a solid metal object will not be entirely supported by water

and so will sink to the bottom, the fact that it is submerged in a fluid will diminish its apparent weight.

Figure 2-18 shows a rectangular block of metal suspended from scales on a string, first in air, and then submerged in a container of some liquid whose density is d gm/cm³. We should not be surprised to notice that the scale reading when the block is submerged is less than the reading when it hangs freely in the air. Apparently the liquid is exerting an upward buoyant force, so that the string has to hold less than the actual weight. Let us figure out what we might expect.



A



B

FIG. 2-18 A block weighed in air (A), and weighed when submerged in a tank filled with liquid of density d (B).

In Fig. 2-18B, there is a downward pressure, equal to hd, on the top of the block. So the total downward force on the top is $F = PA = hd \times ab = hdab$. But there is also an *upward* pressure against the bottom of the block; the bottom is of course submerged deeper than the top, and is $h + c$ cm below the water surface. The upward pressure is thus $(h + c)d$, and the upward force is $(h + c)d \times ab = hdab + cdab$. The buoyancy is the difference between the larger upward force and the smaller downward force:

$$\begin{aligned}
\text{Buoyancy} &= F_{\text{up}} - F_{\text{down}} \\
&= hdab + cdab - hdab \\
&= cdab = d \times abc \\
&= \text{density of fluid} \times \text{volume of fluid displaced} \\
&= \text{weight of fluid displaced.}
\end{aligned}$$

Although we have worked this out for a rectangular block, the result applies to an object of any shape, and leads us to

Archimedes' law: Any body, either wholly or partly submerged in a fluid, is buoyed up with a force equal to the weight of the fluid displaced.

This famous law was discovered by Archimedes, who, as the story goes, thought of it while sitting in a bathtub, and then, in his excitement, rushed wet and naked through the streets of Alexandria, shouting "*Eureka!*" ("I have found it!") The populace of the city was not impressed by the great discovery, however; undoubtedly they thought he had merely found a missing cake of soap in the tub. Whether the above account is true or not, a more credible story is that Archimedes used this law while checking the authenticity of a golden crown that was suspected of being made of a gold-plated cheap metal rather than pure gold. He had to determine its worth without breaking or scratching its surface, so he weighed it when it was suspended on a string in a basin of water and then compared the result with its normal weight. Let us imagine that the crown weighed 2700 gm normally and that when it was submerged in water it appeared to weigh only 2560 gm. The crown's loss of weight was 2700 − 2560 = 140 gm, and was due to the buoyant

effect of the water, which, by Archimedes' law, was just the weight of the water the crown displaced. The density of water is exactly 1 gm/cm^3, so the volume of water displaced by the crown equaled 140 cm^3, which was, of course, also the volume of the crown. Since density equals weight divided by volume, we can figure the density of the crown as 2700/140 $\cong$ 19.3 gm/cm^3, which is just the density of gold, and so Archimedes declared the crown-maker to be an honest man.

Archimedes' law applies also, of course, to floating objects only partially submerged in water. In this case, the weight of a floating object, such as a ship, is the same as the weight of the water it displaces.

We can try another example, a bit more complicated. Suppose a 200-gm stone weighs 140 gm when submerged in water and 150 gm when submerged in oil. Because the stone loses 60 gm in weight when it is in water, it must displace 60 gm of water, and its volume is therefore 60 cm^3. This gives a density of $200/60 = 3.33$ gm/cm^3 for the stone. In the oil, the 60-cm^3 stone loses only 50 gm of weight, so 60 cm^3 of oil must weigh 50 gm, and the density of the oil is $50/60 = 0.83$ gm/cm^3.

Questions

(2-2)

1. A board weighing 35 lb is balanced at its midpoint on a pivot. What upward force does the pivot exert on the board?

2. A steel rod weighing 47 lb is supported at its midpoint by a clamp. What upward force does the clamp exert on the rod?

3. Two children use the board of Question 1 as a seesaw. A girl weighing 60 lb sits 6 ft from the pivot; a 72-lb boy sits 5 ft from the pivot. (a) Is the seesaw still in balance ? (i.e., do the two children exert equal and opposite torques about the pivot?) (b) What total upward force does the pivot exert on the board?

4. A mechanic hangs two buckets on the rod of Question 2. One bucket, weighing 80 lb, is hung 3.5 ft from the clamp; the other bucket is placed 5.0 ft on the other side of the clamp, and weighs 56 lb. (a) Is the rod still in balance on the clamp? (i.e., do the two buckets exert equal and opposite torques about the clamp?) (b) What total upward force does the clamp exert on the rod?

5. Calculate the torque on the plank of Question 3 about the point at which the boy is sitting. Does $\Sigma\tau = 0$ about this point?

6. Calculate the torque on the rod of Question 4 about the point where the 80-lb bucket is hung. Does $\Sigma\tau = 0$ about this point?

(2-3)

7. A tapered pole 12 ft long weighs 48 lb, and balances at a point 4 ft from one end. Two men (one at each end) now lift the pole. How much force does each of the men have to exert?

8. An irregular steel bar 18 ft long weighs 200 lb. Its center of gravity is 8 ft

from one end. A support is placed under each end. What weight is each of the supports held by?

9. Each end of a tapered pole 8 ft long is hung from a spring balance. The balances read 6 lb and 15 lb, respectively. Where is the center of gravity of the pole?

10. At a highway checkpoint, the front wheels of a truck are driven onto platform scales, which read 3600 lb. The truck moves forward so its rear wheels (only) are on the scales, which now read 6200 lb. Where is the center of gravity of the truck? (Its front and rear axles are 12 ft apart.)

11. A man has caught a large fish which he wishes to weigh. He has two spring scales, each able to weigh up to 10 lb, but the fish weighs more than 10 lb. So he takes a light stick 3 ft long, suspends each end of the stick from one of the scales, and hangs the fish on the stick. The two scales now read 6 lb and 8 lb, respectively. (a) What does the fish weigh? (b) At what point on the stick did he hang the fish? (c) If he had hung the fish from a point on the stick just 1 ft from one of the scales, what would each of the scales have read?

12. A tapering bamboo pole is 6 ft long and weighs 10 lb. It is found to balance at a point 2.5 ft from the large end. How much weight must be hung from the small end of the pole to make it balance at its midpoint?

(2-4)

13. Look up the following: sin 23°, tan 87°, cos 60°, tan 30°, sin 60°, cos 43°.

14. Look up the following: cos 19°, sin 58°, tan 27°, sin 37°, tan 11°, cos 71°.

15. A vector 20 units long points in the direction 30° east of north. Calculate trigonometrically its (a) northerly component; (b) its easterly component.

16. As the crow flies, Smithville is 18° south of west from Jonestown, and is 34 miles distant. Smithville is (a) how far south of Jonestown? (b) how far west?

17. Given the following three vectors: $A = 5$ units north, $B = 10$ units 30° east of south, and $C = 10$ units 45° east of north, (a) choose some convenient scale and add the three vectors graphically in at least two different ways, such as $R = A + B + C$ and $R = C + A + B$. (b) Break A, B, and C into their N-S and E-W components, and add these components to find the N-S and E-W components of the resultant.

18. Given three vectors: $\alpha = 10$ units 15° south of east, $\beta = 25$ units 37° east of north, $\gamma = 20$ units 18° west of south; (a) add the vectors graphically in at least two different ways, such as $R = \alpha + \beta + \gamma$ and $R = \beta + \gamma + \alpha$. (b) Break α, β, and γ into their N-S and E-W components, and add these components to find the N-S and E-W components of the resultant.

(2-5)

19. What is the magnitude and direction of R in Question 17?

20. What is the magnitude and direction of R in Question 18?

21. Two tractors are attached to a stump. Because of the terrain, one tractor must pull 45° north of east, the other 60° south of east. Assuming each tractor can exert a pull of 6000 lb, in what direction will the stump move, and what total force will be exerted on it?

22. A charged particle moves in a region where it is influenced by two electric

fields; field *A* exerts on the particle a force of 3.2×10^{-16} newton in a direction 27° east of south, and field *B*, a force of 1.2×10^{-15} newton, directed 38° north of east. What is the total force on the particle and in what direction does it act?

23. A 120-lb weight is suspended from a strut supported by a wire brace, as shown in Fig. Q23. Compute (a) the tension in the wire, and (b) the compressional force on the strut. (The weight of the strut is negligible.)

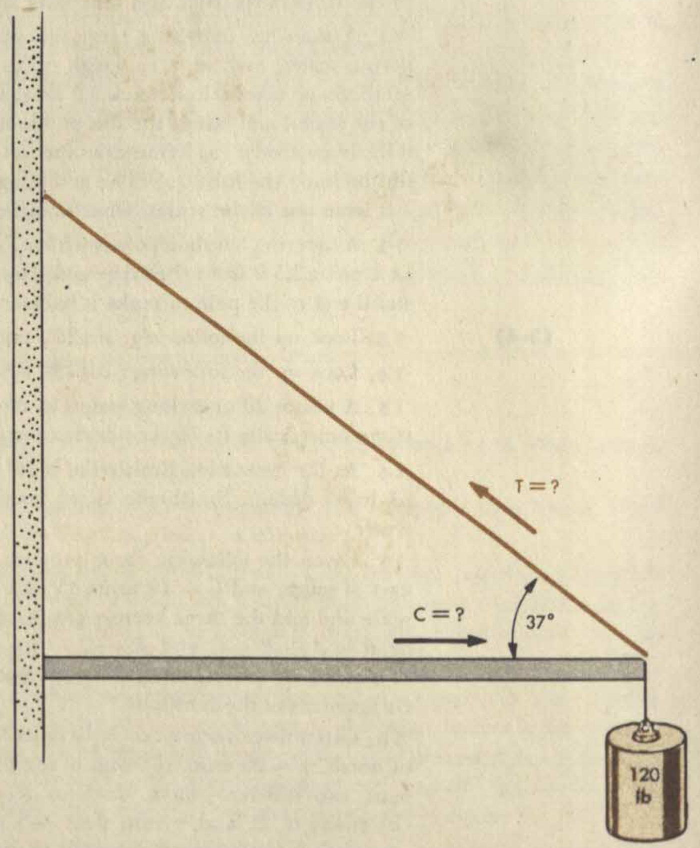


FIG. Q23

24. A horizontal strut of negligible weight is supported by a wire brace, as shown in Fig. Q23. The suspended weight is 200 lb, and the wire brace makes an angle of 30° with the strut. What is (a) the tension in the wire? (b) the compressional force on the strut?

25. A uniform 8-ft strut (whose center of gravity is therefore at its center) weighs 60 lb. It carries an 80-lb load and is supported by a guy wire, as shown in Fig. Q25. (a) What is the tension in the guy wire? (b) What are the horizontal and vertical components of the push of the wall against the strut?

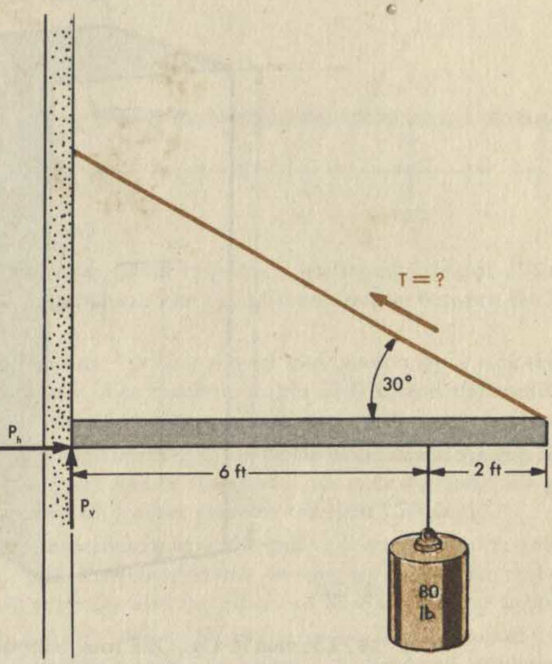


FIG. Q25

26. A uniform, horizontal 150-lb strut is 12 ft long, and is braced with a wire as in Fig. Q25. The wire makes an angle of 42° with the strut. The strut supports a load of 200 lb, suspended 4 ft from its outer end. (a) What is the tension in the wire? (b) What are the horizontal and vertical components of the reaction of the wall against the end of the strut?

27. A large crate weighs 200 lb, and its center of gravity is located as shown in Fig. Q27. How hard must the man pull on the rope in order to lift the back of the crate a short distance off the floor?

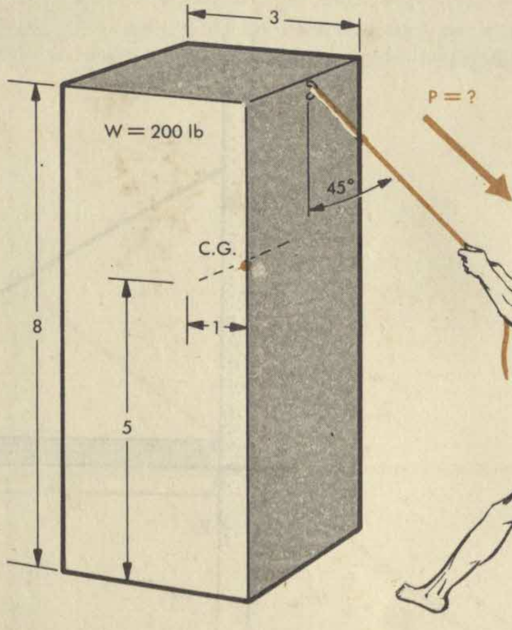


FIG. Q27

28. The man in Fig. Q27 tries to lift the *front* of the same crate a short distance off the floor. He does this by pushing upward at 45°, in a direction opposite to his pull in Question 27, and on the same corner. How hard must he push?

29. A body is acted on by three forces, as shown in Fig. Q29. The body is not in equilibrium, but it is possible to add one force which will hold it in equilibrium. What is the magnitude and direction of this force, and where on the body must it be applied?

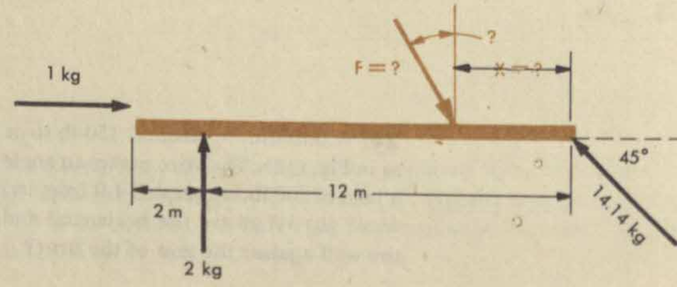


FIG. Q29

30. A body is acted on by three forces, as shown in Fig. Q30. The body is not in equilibrium, but it is possible to add one force which will hold it in equilibrium. What is the magnitude and direction of this force, and where on the body must it be applied?

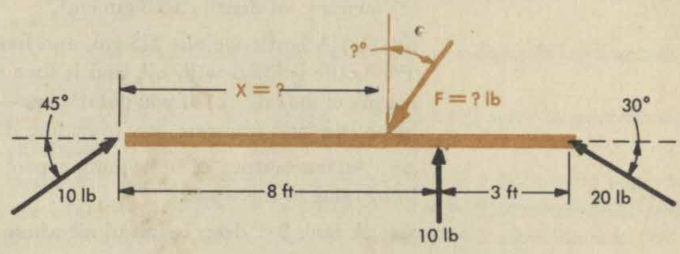


FIG. Q30

(2-6) **31.** A box weighing 120 lb requires a horizontal force of 25 lb to drag it across a level floor. What is the coefficient of friction between the box and the floor?

32. A heavy machine is being shoved into position by a jack that applies a horizontal force to it. The machine weighs 7000 lb, and the coefficient of friction between its base and the floor is 0.35. How much force must the jack exert?

33. In the experiment bracketing the center of gravity of the golf club between fingers, would the outcome be the same if one wore a smooth silk glove on the left hand, and a rough leather glove on the right? Explain.

34. Would the experiment with the golf club on the fingers have the same outcome if the shaft were covered with fine sandpaper from the center of gravity to the end of the handle and the other end were smoothly polished? Explain.

35. In order to frisk a suspect, the police have him lean forward with his arms horizontal and his hands against a smooth wall. His hands are greasy, so there is no friction between hands and wall. The coefficient of friction between his shoes and the floor is 0.40, and his center of gravity is midway between his shoulders and his feet. What is the minimum angle θ between the suspect's straight body and the floor that can be allowed without the suspect's falling?

36. A man stands at the midpoint of a 20-ft ladder which leans against a smooth vertical wall. The coefficient of friction between the bottom of the ladder and the floor is 0.20. How far out from the base of the wall may the foot of the ladder be placed, if it is not to slip? (Neglect weight of ladder.)

37. A reluctant 120-lb dog ($\mu = 0.50$) is pulled along a path by a rope making an angle of 40° with the horizontal. How strong a pull is required?

38. A man slides a 200-lb box along a floor ($\mu = 0.30$) at constant speed, with a push directed 37° down from the horizontal. How hard must he push?

(2-7) **39.** Knowing that water has a density very close to 1 gm/cm³, compute the density of water in lb/ft³.

40. Knowing the density of water in gm/cm³ and in lb/ft³ (see Question 39), calculate the density, in lb/ft³, of the oil and mercury mentioned in Question 43.

41. The density of copper is 8.90 gm/cm³. What is the mass of a block of copper 4.0 cm × 5.0 cm × 10.0 cm?

42. A piece of iron 6.0 cm $\times$ 12.0 cm $\times$ 15.0 cm has a mass of 8.42 kg. What is the density of the iron, in gm/cm^3?

43. A volumetric flask holds exactly 200.00 cm^3 and weighs 230 gm, empty. What would it weigh when filled with (a) water? (b) oil, of density 0.82 gm/cm^3? (c) mercury, of density 13.6 gm/cm^3?

44. (a) A bottle weighs 213 gm, and has an internal volume of just 800 cm^3. The bottle is filled with oil, and is then found to weigh 861 gm. What is the density of the oil? (b) If you did not know the volume of the bottle, what would be an easy and accurate way of finding it?

45. At the bottom of a swimming pool 10 ft deep, what is the pressure (in lb/in^2) due to the water?

46. A tank 8 ft deep is full of oil whose density is 0.90 gm/cm^3. What is the pressure (in lb/in^2) on the bottom due to the oil?

47. The swimming pool of Question 45 is 12 ft $\times$ 20 ft. What is the total force on its bottom?

48. The tank of Question 46 is 6 ft in diameter. What is the total force on its bottom?

49. A tank 5 m high is half filled with water, and then is filled to the top with oil of density 0.85 gm/cm^3. What is the pressure at the bottom of the tank due to these liquids in it?

50. The bottom 15 cm of a vertical glass tube 125 cm long is filled with mercury; the tube is then filled to the top with carbon tetrachloride (density 1.59 gm/cm^3). What is the pressure at the bottom of the tube?

(2-8)

51. A hydraulic jack has a piston 0.5 cm in diameter on which a force is applied, and a piston 6 cm in diameter which raises a load. How much force must be applied to lift a load of 2000 kg?

52. In a hydraulic jack, the piston that raises the load is 3 inches in diameter; the operating force is applied to a piston 0.5 inch in diameter. How many pounds of force must be applied to raise a load of 6 tons?

(2-9)

53. A stone weighing 250 gm appears to weigh only 150 gm when submerged in water. (a) What is the loss of weight? (b) What is the net upward buoyant force on the stone? (c) What weight of water does the stone displace? (d) What is the volume of this amount of water? (e) What is the volume of the stone? (f) What is the density of the stone?

54. A stone weighing 750 gm appears to weigh only 500 gm when it is submerged in water. (a) What is the volume of the stone? (b) its density?

55. A metal block weighing 3000 gm appears to weigh only 2600 gm when submerged in oil of density 0.80 gm/cm^3. (a) What is the volume of the block? (b) What is its density?

56. A 1520-gm metal bolt appears to weigh 1202 gm when submerged in carbon tetrachloride (see Question 50). (a) What is the volume of the bolt? (b) What is its density?

57. In order to find the density of some acid, we first weigh a glass stopper in air (250 gm), then in water (150 gm), and then in the acid (125 gm). (a)

What is the volume of the stopper? (b) What is the density of the stopper? (c) What is the density of the acid?

58. In order to find the density of some oil, a piece of metal is weighed in air (300 gm), then in water (200 gm), then in the oil (220 gm). (a) What is the volume of the piece of metal? (b) What is the density of the metal? (c) What is the density of the oil?

59. A block of wood (density 0.60 gm/cm^3) floats on water. What fraction of the volume of the block is submerged?

60. A block of wood floats on oil (density 0.80 gm/cm^3) with 60 percent of its volume submerged. What is the density of the wood?

61. If a boat loaded with rocks floats in the middle of a swimming pool, and a man in the boat throws the rocks overboard, what will happen to the water level of the pool? Explain. (At a scientific meeting, this question was put to Dr. Gamow, the physicist J. Robert Oppenheimer, and Nobel prize-winner Felix Bloch. All three of them, not thinking too carefully, gave the wrong answer!)

62. A boat loaded with logs of light wood (density less than 1 gm/cm^3) floats in a swimming pool. A man in the boat throws the logs overboard. What happens to the water level of the pool? Explain.

63. Air is a fluid (as all gases are), and Archimedes' principle applies to bodies in air, as well as to those in the denser liquids. Ordinarily the buoyancy of air is so small that it can be neglected, but suppose a balloon with a volume of 3000 m^3 is filled with helium, which has a density of 1.8×10^{-4} gm/cm^3. (a) Assuming air has a density of 1.3×10^{-3} gm/cm^3, calculate the pull on the mooring rope if the balloon, with all its equipment, weighs 1600 kg. (b) What would the pull be if the balloon were filled with hydrogen instead of helium? (The density of hydrogen is about 9×10^{-5} gm/cm^3.)

64. A man needs a very precise value for the weight of a plastic cube 10 cm on a side. On an accurate balance, he balances the plastic cube on one side against 2980.2 grams of brass weights (density 8.6000 gm/cm^3) on the other side. What is the exact weight of the plastic? (See Question 63 for other data.)

chapter three

Bodies in Motion

3-1 The Measurement of Motion

In order to investigate the behavior of moving bodies, we must first analyze the general idea of motion itself. What properties or characteristics of motion can we measure or calculate to help us describe the behavior of a moving body? The *distance* traveled and the *time* required to travel the distance are two obvious factors, and the idea of *average speed* is one that our own traveling has made familiar to all of us: if it takes 5 hours to drive to a city 200 miles away, our average speed is 200 miles/5 hours = 40 mi/hr. Without thinking of it, we have made use of the relationship

$$v_{av} = \frac{d}{t}$$

where the letter v stands for speed. Multiplying both sides of this equation by t, we get

$$d = v_{av}t.$$

If we average 30 mi/hr, in 4 hr we travel a distance of 30 mi/hr × 4 hr = 120 mi.

It is obvious that v_{av} alone tells us nothing about the speed at any instant during the time interval we have considered. A car slows, stops, and speeds up again many times during a four-hour trip, and we must have some way of taking into account the way in which the speed of a body changes if we want to describe its motion properly. In order to do this, we make use of the idea of *acceleration*, which is defined as *the rate of change of velocity*.

So far, we have been talking about the *speed* of a moving body—why do we instead use the word *velocity* when we define acceleration? In ordinary conversation, we generally use "speed" and "velocity" interchangeably, but the physicist gives somewhat different meanings to the two words. *Speed refers only to the rate at which distance is covered, without regard for direction; velocity includes both speed and direction.* A car going north at 40 mi/hr has the same *speed* as a car going east at 40 mi/hr, but their *velocities* are different, because their directions are not the same. Later on, when we discuss rotating bodies and bodies moving in curved paths, we shall see that a change in direction represents as real an acceleration as a change in speed, and for this reason, acceleration is defined as the rate of change of *velocity*, which includes both. For the present, however, we shall not consider changes in direction, so that our velocity changes will be only changes in speed.

If we step on the gas pedal of a car, the speed of the car will change—for example, from 5 mi/hr to 40 mi/hr in 10 sec. Put into mathematical form, our definition of acceleration is

$$a = \frac{\Delta v}{t}.$$

(The Greek capital letter Δ—*delta*—is the mathematician's abbreviation for "the change in." The change will of course be positive if the quantity is increasing, and negative if it is decreasing.)

Substituting our figures into this equation, we get

$$a = \frac{40 \text{ mi/hr} - 5 \text{ mi/hr}}{10 \text{ sec}} = \frac{35 \text{ mi/hr}}{10 \text{ sec}} = 3.5 \text{ mi/hr/sec}$$

which is to say that in each second, the speed has increased by 3.5 mi/hr. These are somewhat clumsy (although perfectly proper) units, and acceleration is more often given in ft/sec/sec, abbreviated ft/sec^2 (or cm/sec^2, or m/sec^2).

If a body starts from rest, moving with a constant acceleration of a ft/sec^2, at the end of 1 sec it will have a velocity of a ft/sec; at the end of 2 sec its velocity is $2a$ ft/sec, and so on. We can write this in a general form:

$$v_t = at$$

where v_t is the velocity at the end of t sec. This is a very useful equation

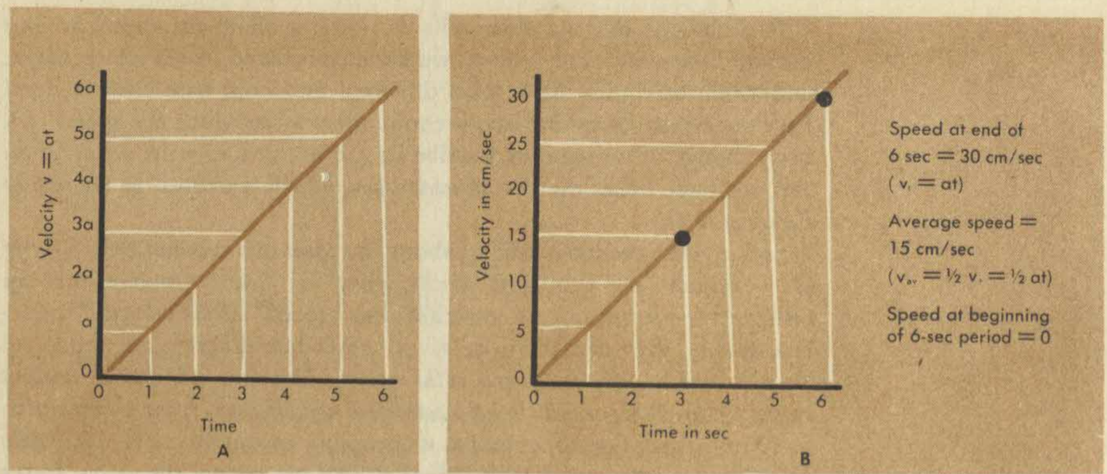


FIG. 3–1 Time vs. velocity for uniformly accelerated motion.

for some purposes, but it is often convenient to have a relationship which includes the distance the body has moved from its starting place. To consider how to do this, look at Fig. 3-1A, in which the velocity (or speed) has been plotted against time for a uniformly accelerated motion—that is, motion with a constant, unchanging acceleration a. Since the increase in velocity is the same for each second, the graph is a straight line, and the *average* velocity v_{av} during any time interval is equal to the velocity at the middle of the interval. For example, the average velocity during the interval from $t = 2$ to $t = 4$ is $3a$, the velocity that corresponds to t when $t = 3$. If we start from rest at $t = 0$, the average speed of the body during the next t seconds will be $\frac{1}{2}at$. Figure 3-1B shows a numerical example.

To get the distance d that the body will have moved, we need only multiply this *average* speed by the time. (A car averaging 40 mi/hr will go 200 mi in 5 hr.) We can now write

$$d = v_{av} \times \text{time} = \tfrac{1}{2}at \times t = \tfrac{1}{2}at^2.$$

By eliminating the time from the two equations we already have, we can get a third relationship. We know $v_t = at$, and solving for t we get $t = v_t/a$, so that $t^2 = v_t^2/a^2$. Now we can substitute this value for t^2 in the equation $d = \frac{1}{2}at^2$:

$$d = \frac{1}{2} \times a \times \frac{v_t^2}{a^2} = \frac{v_t^2}{2a}$$

and

$$v_t^2 = 2ad.$$

So far, we have developed equations of motion that can be used *only when the accelerated motion starts from rest.* These will be much more useful if we can extend them to include bodies which are originally moving with a velocity we can call v_o.

The equation $v_t = at$ tells us how much the speed has increased because of the acceleration. This amount needs only to be added to the original speed, and we have

$$v_t = v_o + at$$

for this more general case.

The next equation, $d = \frac{1}{2}at^2$, gives the distance covered because of the acceleration; if the body were originally moving at the speed v_o, in t seconds it would have gone a distance v_ot without any acceleration. By adding these together, we get

$$d = v_ot + \tfrac{1}{2}at^2.$$

In a similar manner, our last equation becomes

$$v_t^2 = v_o^2 + 2ad.$$

Let us tabulate these equations, with a comment on each:

1. $v_t = v_o + at$ (does not include the distance)
2. $d = v_ot + \frac{1}{2}at^2$ (does not include the final velocity)
3. $v_t^2 = v_o^2 + 2ad$ (does not include the time)

Which one (or ones) we want to use will depend on what data we have given in a problem and what answer we are trying to find. For example, what is the acceleration of a car that goes from a speed of 20 ft/sec to 80 ft/sec in 15 sec? We know v_o, v_t, and t; we want to find a. The first equation includes these items, and we write

$$80 = 20 + 15a$$

$$15a = 80 - 20 = 60$$

$$a = \frac{60}{15} = 4\ \text{ft/sec}^2.$$

We might now want to find out how far the car moves during this period of acceleration. Since the acceleration is now known, we can use the third equation:

$$(80)^2 = (20)^2 + 2 \times 4 \times d$$

$$8d = 6400 - 400 = 6000$$

$$d = \frac{6000}{8} = 750\ \text{ft}.$$

(Can you get this same answer in another way?)

In this example, the acceleration, the velocity, and the distance have

all been in the same direction. Often, however, this is not the case, and we must carefully watch + and − signs. Consider a car traveling north at 30 m/sec. The driver applies the brake and reduces his speed by 5 m/sec each second. This rate of reducing speed is *deceleration*, which subtracts from his speed rather than adds to it. It is actually acceleration toward the south. If we choose + to indicate vectors pointing north, the acceleration will be negative: $a = -5\ \text{m/sec}^2$. To find how long it will take him to stop, we may use our original equation

$$v_t = v_o + at.$$

Since v_t, the final speed, is zero, we get

$$0 = +30 - 5t$$

or

$$t = 6\ \text{sec}.$$

If we had chosen south to be our positive direction, we would have written

$$0 = -30 + 5t$$

and

$$t = 6\ \text{sec}.$$

3-2 The Cause of Motion

Even a baby soon learns that there is some sort of a direct connection between force and motion—a push or a pull will often cause things to move. However, exactly *how* force and motion are related is not nearly so obvious. In fact, the basic ideas concerning motion itself were not really investigated until the beginning of the seventeenth century. Then the Italian physicist Galileo analyzed the relationships between distance, speed, and time, somewhat as we have done in the preceding section; he was probably the first man to understand the concept of acceleration and to appreciate its importance in the study of motion.

The year Galileo died in Italy (1642), the great physicist and mathematician Isaac Newton was born in England. Newton built on the foundations so well laid by Galileo, and his further studies led him to three fundamental laws of motion.

Newton's first law of motion: a stationary body will remain motionless, and a moving body will continue to move in the same direction with unchanging speed unless it is acted on by some unbalanced force.

This law is an expression of the property of matter called *inertia*, which describes its resistance to having its velocity changed in any way. The law also recognizes that the velocity of a body *can* be changed,

but only by the application of a net force. Let us return to the book on the desk, and give it a slight sidewise push. If the push is a gentle one, the book does not move. We know why it does not; our push is opposed, or "balanced" by the equal and opposite force of friction. Only if our push exceeds the opposing force of friction does the book begin to move.

The situation is less obvious when we push the book across the desk with constant speed in a straight line—in other words, with an unchanging velocity. In this case, also, the force of friction exactly balances, or opposes, the force of the push, so that the net force is zero. If we push harder, we exceed the force of friction and the book speeds up; if we stop pushing, the unopposed force of friction quickly brings the book to a halt.

All this is well and good, but it does not bring us to grips with the main problem: exactly what happens when the forces on a body do *not* add up vectorially to zero? For many centuries, scientists and philosophers had missed the point of this question, because they tried in some uncertain way to relate the force and the *velocity* which it presumably caused. Galileo almost grasped the point, and Newton grasped it clearly; the direct relationship is not between force and velocity, but between force and *acceleration.*

Newton's second law of motion: The acceleration given to a body by a force applied to it is directly proportional to the force, and is in the same direction as the force. If the same force is applied to bodies of different masses, the accelerations produced will be inversely proportional to their masses.

3-3 Mass and Weight

Before we put Newton's second law into the form of an equation, we should pause and give some thought to the meaning of "mass," which is such an important part of the law. Matter of any kind has two universal characteristics: it is pulled by the force of gravity—that is, it has *weight*—and it resists being accelerated—that is, it has *inertia.* We can use either of these characteristics to measure the mass of a body.

Weight is something that can be measured directly by means of a spring balance. For example, let us take a standard kilogram (a piece of metal that the manufacturer has made to be an accurate duplicate of the standard kilogram in the vault at Sèvres) and hang it from a spring balance. Before we mark the position of the indicator, however, we should pause and consider. Suppose we make a mark when we are in Boulder, Colorado; then we go to San Francisco—here (because we are at a lower altitude and are hence closer to the center of the earth) the pull of gravity is stronger, the kilogram mass actually weighs more, and the mark will be in a slightly different place. Conversely, if we take

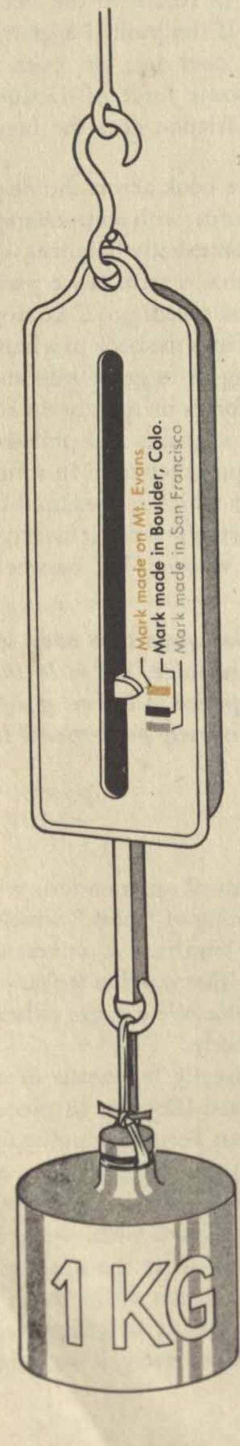


FIG. 3-2 A kilogram weighs different amounts in different locations.

the spring balance and standard kilogram to the high peak of Mount Evans in the Colorado Rockies, the weight is less, and we shall have to make a third mark (Fig. 3-2).

Which of these three marks should we choose to be permanently engraved? What does the manufacturer of spring balances do? Since most of the population lives not far above sea level, he probably marks it a little above the San Francisco mark and hopes for the best. The ordinary spring balance is rather crude and does not pretend to any great accuracy. If it is a little off for the customers in Death Valley or in Boulder, no harm is done, as the user does not expect precision anyway.

The point of this discussion is that the kilogram and the gram are units of *mass* and that their weights vary from place to place. If we use another, and more common, sort of balance—a platform or an analytical balance (Fig. 3-3)—we find that this confusion disappears. Select a stone that will exactly balance the kilogram in Boulder. It will also exactly balance the kilogram in San Francisco or on Mount Evans, because, although the pull of gravity varies from place to place, it remains everywhere equal for the kilogram and the stone. What we can say from "weighing" of this sort is that the mass of the stone is the same as the mass of the kilogram—or, more briefly, that the *mass* of the stone is one kilogram. What the actual *weight* of the stone is we do not know and cannot compute until we have some way of taking into account exactly how strong the pull of gravity is at the particular location we are interested in.

In modern commerce and engineering, the pound (unlike the gram and the kilogram) is *not* used as a unit of mass. Instead, its weight at 45° latitude and at sea level is used as a unit of *force*. With this definition in mind, we can successfully calibrate our spring balance. Take a standard pound to Calais, Maine, or Portland, Oregon (both of these cities nearly fulfill the requirements for latitude and altitude), hang it on the balance, and you can confidently mark the indicator position as "1 lb." Now if we take our apparatus to Boulder or Mount Evans, we need not be surprised to find that a standard pound does not weigh a pound. Actually, the pull of gravity is smaller, and its weight *is* less than a pound.

Suppose we were set down in some unknown location and were handed a stone from the roadside. With a platform balance and a set of calibrated masses (incorrectly called "weights," a misnomer we must learn to live with), we could accurately determine the mass of the stone in grams or kilograms but would be unable to measure its weight. On

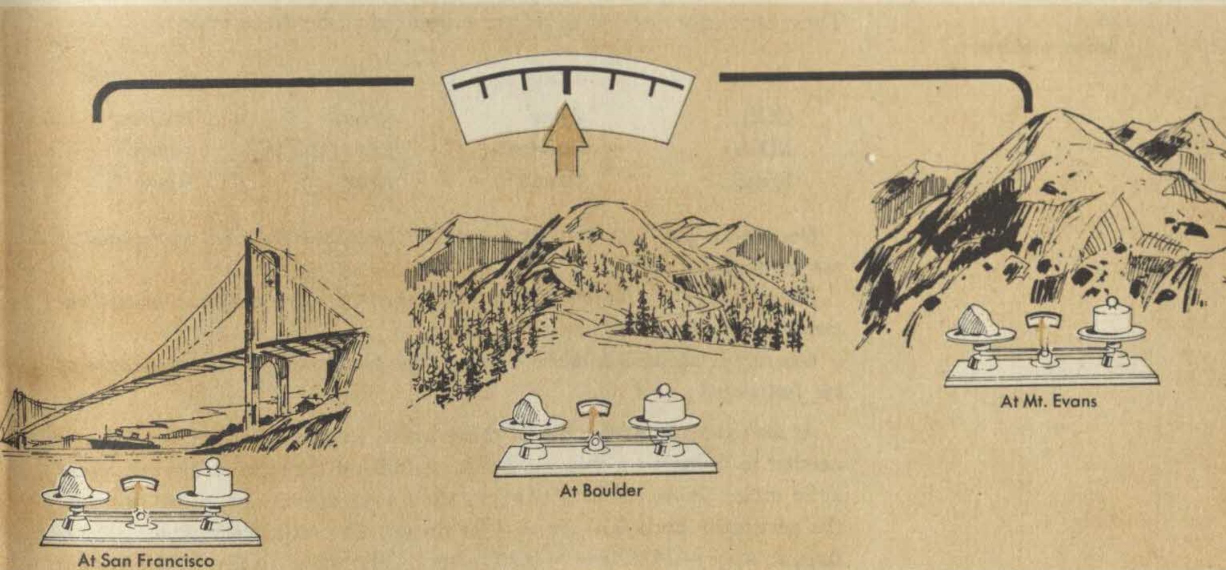


FIG. 3–3 The mass of an object, determined by comparison with a standard mass, is the same in all locations.

the other hand, an accurately calibrated spring balance would tell us that the weight of the stone was exactly so many pounds but would leave us ignorant about its mass. We need some procedure for relating mass and weight, and for this we must turn to the other property all material objects have—inertia.

Newton's second law states that (for equal masses) acceleration is proportional to force; and (for equal forces) acceleration is inversely proportional to mass. We can write this algebraically as

$$a = k \times \frac{F}{m}.$$

We have defined length and time units with which to measure a, and mass units (the gram and the kilogram) with which to measure m in the metric systems. We have a force unit (the pound) in the British system. So far we have no force unit in the metric systems, and no mass unit in the British system. It would seem to be very convenient if these missing units were defined in such a way as to make the proportionality constant k equal to 1 in Newton's second law. It seemed so, too, to the nineteenth-century scientists who then defined them. When we use these units, we may write $a = F/m$, or more usually

$$F = ma.$$

These specially defined units are indicated in boldface type:

System	*Force*	*Mass*	*Acceleration*
CGS	**dyne**	gram	cm/sec^2
MKS	**newton**	kilogram	m/sec^2
British	pound	**slug**	ft/sec^2

One dyne is the force necessary to give a mass of one gram an acceleration of one centimeter/second².

One newton (nt) is the force necessary to give a mass of one kilogram an acceleration of one meter/second².

One slug is the mass to which a force of one pound will give an acceleration of one foot/second².

As an example of the use of these units, let us find what net force is needed to bring to a stop, in 2 min, a 200-ton (metric) train traveling at 24 m/sec. If we arbitrarily select the forward direction to be positive, the necessary Δv is $0 - 24 = -24$ m/sec. The required acceleration is $\Delta v/t$, so $a = -24/120 = -0.20\ m/sec^2$. The mass (in the MKS system we are using) must be in kilograms; $m = 200{,}000 = 2 \times 10^5$ kg. Then

$$F = ma$$
$$= 2 \times 10^5 \times (-0.20) = -4 \times 10^4 \text{ nt}.$$

The negative sign on the force in the answer shows it to be toward the rear—opposite to the direction we selected to be positive.

Or, consider a car whose mass is 100 slugs. If it can exert a braking force of 1600 lb, how many feet will it travel in coming to a stop from a speed of 60 ft/sec? We can easily calculate the distance if we know the acceleration; and since we know both the mass and the force, the acceleration follows simply from Newton's second law:

$$F = ma\ ; \quad a = \frac{F}{m} = \frac{1600}{100} = 16 \text{ ft/sec}^2.$$

Then

$$v^2 = v_0^2 + 2ad$$
$$0 = (60)^2 + 2 \times 16 \times d\ ; \quad d = \frac{-3600}{32} - 113 \text{ ft}.$$

(We have implicitly made the acceleration of braking positive when we made the force that caused it positive in the equation above. This put the positive direction toward the rear of the car, and the -113 ft thus indicates that the car will travel *forward* 113 ft before it stops.)

Again, we might deal with an electric field that exerts a force of 1.8×10^{-10} dyne on an electron, whose mass is 9×10^{-28} gm. Under these conditions, what is the acceleration of the electron?

$$F = ma\ ; \quad a = \frac{F}{m} = \frac{1.8 \times 10^{-10}}{9 \times 10^{-28}} = 2 \times 10^{17} \text{ cm/sec}^2.$$

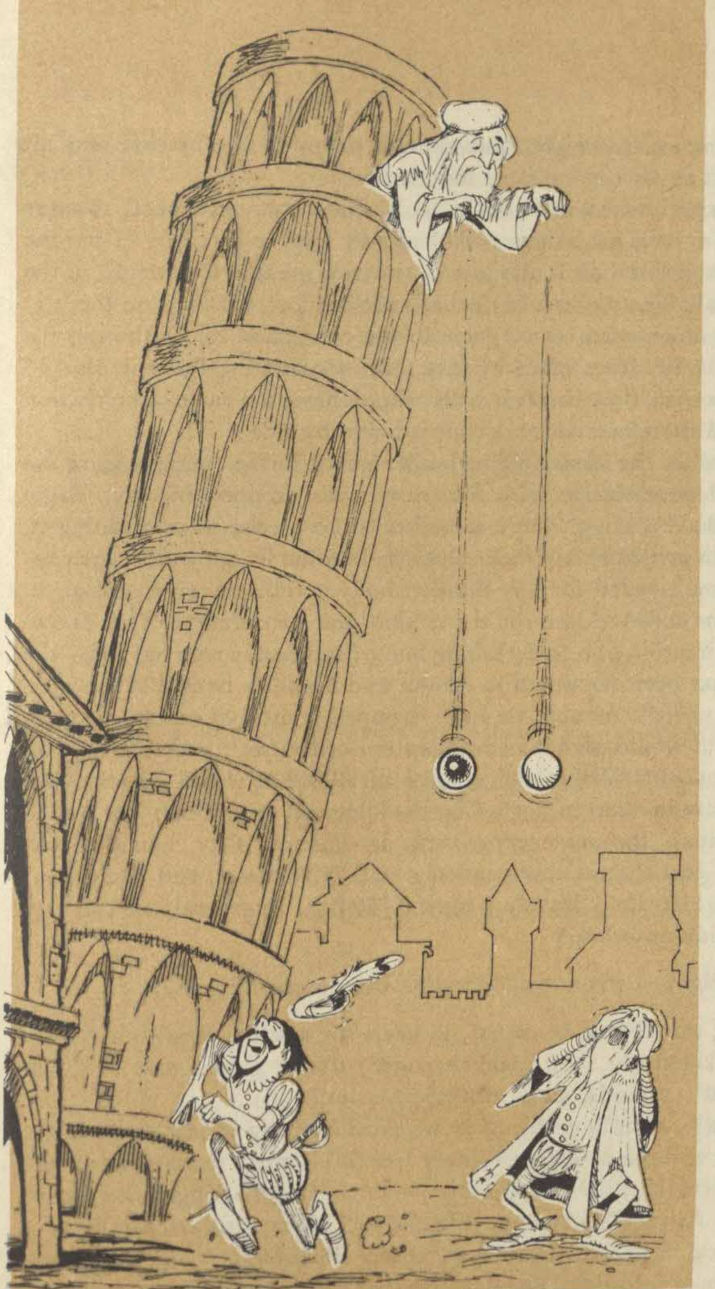


FIG. 3-4 Galileo's experiment with falling bodies.

3-4 Weight and Falling Bodies

The Leaning Tower of Pisa, besides being one of the architectural wonders of the world, is also inseparably connected with the history of physics because of the part it played in an experiment that was alleged to have been performed more than three centuries ago by the famous Italian scientist Galileo. From the upper platform of the Tower, the story has it, Galileo simultaneously released two spheres, a heavy one made of iron and a lighter one made of wood (Fig. 3-4). In spite of a

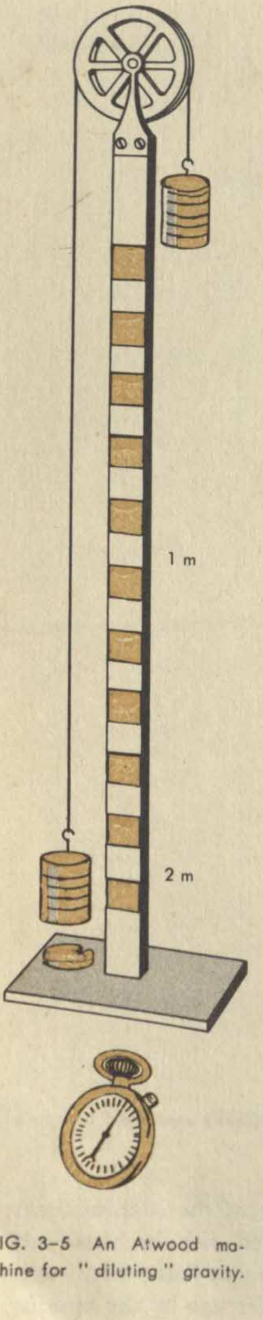


FIG. 3-5 An Atwood machine for "diluting" gravity.

large difference in weight, both spheres dropped side by side and hit the ground at almost the same moment.

Is this experimental result of Galileo's what we should expect? Assume that the two balls have masses of 10 kg and 1 kg, respectively. Then the weight of the iron ball is also just 10 times as great as the weight of the wooden ball. The weights of the balls are the forces that cause them to accelerate downward, so we come to the conclusion that although the iron ball is 10 times more massive, the accelerating force is also 10 times larger, so that the two balls would have the same acceleration and would therefore travel downward side by side.

Let us follow the idea of falling bodies a little further, and measure the downward acceleration of a kilogram mass dropped in, say, Eagle City, Alaska. Although there are other more precise ways of doing it, we could in principle determine this acceleration by accurately measuring the time needed for any massive body to fall any given distance. Then, from $d = \frac{1}{2}at^2$, we could calculate the acceleration. This downward acceleration of a freely falling body (universally referred to by the letter g) has been measured in Alaska and found to be 982.18 cm/sec^2, or 9.8218 m/sec^2. Assume we have dropped a mass of exactly 1 kg—a force of 1 nt would give it an acceleration of 1 m/sec^2; and to give it an acceleration of 9.8218 m/sec^2, a force of 9.8218 nt would be required. We therefore see that in Eagle City the kilogram must weigh just 9.8218 nt. Repeating the same experiment in Panama City, Canal Zone, we would find that at this location $g = 9.7824$ m/sec^2, and could conclude that here the kilogram weighs 9.7824 nt. In general, we can say, for any body anywhere,

$$\text{weight} = \text{mass} \times \text{acceleration due to gravity} = m \times g.$$

(We must remember, of course, to keep the systems of units straight: newtons, kilograms, and meters/second^2; dynes, grams, and centimeters/second^2; pounds, slugs, and feet/second^2.)

The study of falling objects is an excellent way to learn about the basic laws of motion. Since ordinary free fall is much too fast for convenient observation, it is helpful to slow the fall with an ingenious device known as Atwood's machine (Fig. 3-5). This consists essentially of a long vertical support, with a light, nearly frictionless pulley on top and a collection of calibrated metal masses that can be attached in any desired quantities to the ends of the string that runs over the pulley. With equipment as simple or as fancy as our resources allow (an ordinary two-meter stick and a stopwatch will give reasonably good results), we can measure d and t, and so investigate the relationship between force, mass, and acceleration.

Consider an Atwood machine on which we have placed exactly 900 gm on one side, and 800 gm on the other. Obviously the 900-gm

side will start to descend as soon as we release it. What will be its downward acceleration? If 800 gm were on both sides, the weights would be balanced, and there would be no acceleration; the accelerating force is the unbalanced weight of the extra 100 gm on the heavy side.

Our trouble now is that we do not know just what 100 gm weighs. If this book were to be used only in Eagle City, Alaska, we could say at once that 100 gm has a weight of mg dynes $= 100 \times 982.18 = 98{,}218$ dynes. But in the Canal Zone, 100 gm weighs only 97,824 dynes, and so on. Let us reconcile ourselves to an error of a few tenths of a percent and assume that $g = 980\ \text{cm/sec}^2$ or $9.8\ \text{m/sec}^2$ or $32\ \text{ft/sec}^2$ everywhere on the surface of the earth.

So, for the Atwood machine the accelerating force is $100 \times 980 = 98{,}000$ dynes, or 0.98 nt. This force serves to accelerate *both* masses—a total of $900 + 800 = 1700$ gm, or 1.700 kg, and

$$a = \frac{F}{m} = \frac{98{,}000}{1700} = 57.6\ \text{cm/sec}^2$$

or

$$\frac{0.980}{1.700} = 0.576\ \text{m/sec}^2.$$

With this knowledge of a, we can go on to determine distances, times, speeds, etc., in any way that is required.

Let us ask another question about the Atwood machine as it was described above. What is the tension in the string? If we prevent any motion by holding the 900-gm mass, the string will have to support only

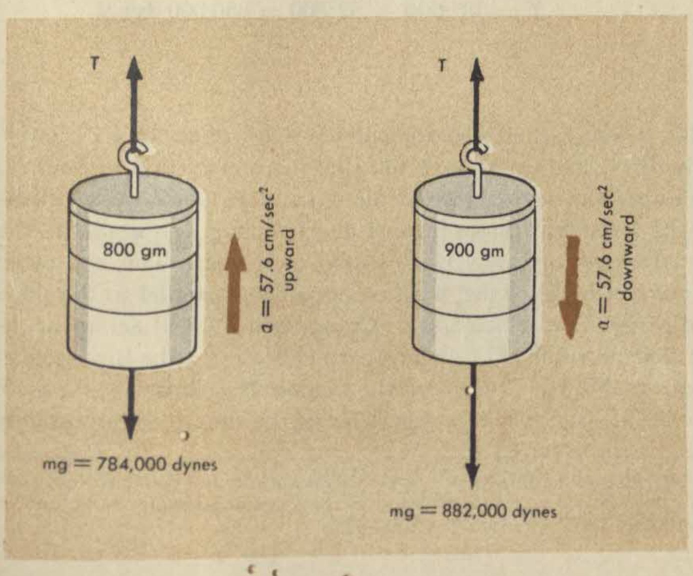


FIG. 3-6 Isolating each of the "weights" on the Atwood machine.

the weight of the 800-gm mass, which will cause a tension of $800 \times 980 = 784{,}000$ dynes. However, if we hold the 800-gm mass, the tension in the string will be the weight of the 900-gm mass, or 882,000 dynes. Is it safe to assume that when the masses on the machine are accelerating freely, the tension will be somewhere between these two values? We can solve this problem by concentrating on one of the masses—the 800-gram one, say—and putting the rest of the machine entirely out of mind. (This is called "isolating" one of the factors of a complex problem. See Fig. 3-6.)

We see only a mass of 800 gm which has an upward acceleration of 57.6 cm/sec². The upward acceleration must be caused by a net upward force of $F = ma = 800 \times 57.6 = 46{,}080$ dynes. This net force is the resultant of the two forces acting on the mass: (1) the tension in the string, which is T dynes upward, and (2) the weight of the mass, which is 784,000 dynes downward. Since T must obviously be greater than the weight, we can write

$$T - mg = 46{,}080$$
$$T = 46{,}080 + mg$$
$$T = 46{,}080 + 784{,}000 = 830{,}000 \text{ dynes.*}$$

This answer can be checked by isolating the 900-gm mass in the same way. Since it is accelerating downward, its weight is greater than the tension T, and we have

$$mg - T = ma$$
$$900 \times 980 - T = 900 \times 57.6$$
$$T = 900 \times 980 - 900 \times 57.6$$
$$T = 882{,}000 - 51{,}800 = 830{,}000 \text{ dynes.}$$

3-5 Inclined Planes

In Fig. 3-7, a boy is shown coasting down a 20° slope. If the snow is reasonably dirty, we may assume the coefficient of friction to be 0.25. In this example, the motion must be along the slope: the boy's acceleration and the force that causes it must also be along, or parallel to the slope. But the pull of gravity W is stubbornly vertical. This can easily be taken care of by resolving W into components parallel to the slope (F_{par}) and perpendicular to it (F_{perp}). Component F_{par} will act to accelerate the sled downhill; F_{perp}, being perpendicular to the direction of motion, can itself have no effect on the motion. It is, however, the force normal to the surface, which presses sled and snow together; it is exactly

* Since our value of g (980 cm/sec²) is somewhat uncertain in the third figure, we are certainly not justified in retaining more than three significant figures in our answer. It has accordingly been rounded off to 830,000.

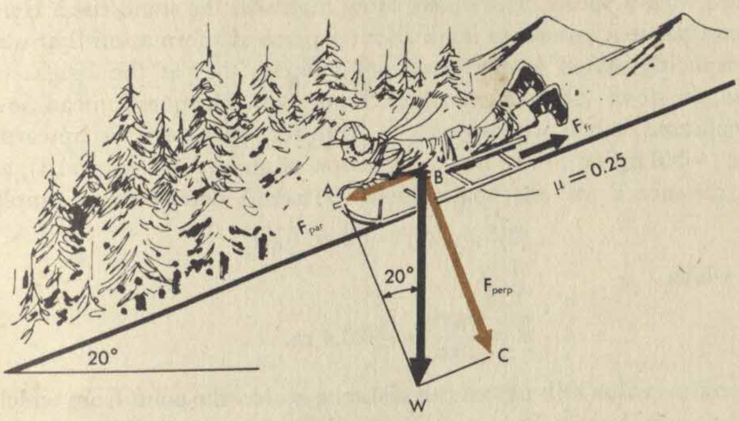


FIG. 3-7 Resolution of the pull of gravity into components perpendicular and parallel to the direction of motion.

what we need in order to figure the retarding force of friction (F_{fr}). So let us calculate the net force accelerating the sled downhill:

$$\begin{aligned} F_{\text{net}} &= F_{\text{par}} - F_{\text{fr}} = F_{\text{par}} - \mu F_{\text{perp}} \\ &= W \sin 20° - 0.25 \times W \cos 20° \\ &= W \times 0.342 - 0.25 \times W \times 0.940 = 0.107\,W \end{aligned}$$

The mass of the sled is not given; it can be assumed to be m in any system we choose. The weight W will then be mg in this same system, and we have

$$F = ma$$

$$0.107\,mg = ma.$$

The mass m cancels out, showing that the results would be the same for a sled of any mass, and then

$$a = 0.107\,g.$$

The same result could be reached by reasoning that since a force equal to the weight of a body gives it an acceleration g, then a force of 0.107 times its weight would give it an acceleration of 0.107 g, in any system of units.

3-6 Projectile Motion

Because all unsupported bodies have a constant downward acceleration produced by the pull of gravity, the equations of uniformly accelerated motion can be applied to falling bodies, to rocks thrown into the air, or to golf balls and bullets. As a simple case, consider a boy who stands

on a bridge 10 m above the water and throws a stone nearly vertically upward with a speed of 20 m/sec. How high will the stone rise? Here we must pause a minute to think about a piece of information that was not explicitly stated in the problem—namely, that at the height of its rise, the stone will momentarily be motionless. The question can now be rephrased: since we know v_0 (+20 m/sec, because v_0 is upward) and a (-9.8 m/sec^2, since the acceleration of gravity is downward), at what distance d will the final velocity be zero? This becomes simply

$$(0)^2 = (20)^2 - 2 \times 9.8d$$

from which

$$d = \frac{400}{19.6} = +20.4 \text{ m}.$$

The positive value tells us that this distance is *above* the point from which the stone was thrown, because we have chosen the upward direction to be +.

We might further ask how long it will be before the stone hits the water. When it does, it will be 10 m below the bridge, so that d will then be -10. (The distance d does *not* represent the total distance traveled, but is the distance of the stone from the point of beginning; as it passes the boy on the way down, d is at that instant zero.) For this calculation, we may use the second equation and get

$$-10 = 20t - \tfrac{1}{2} \times 9.8t^2$$

or

$$4.9t^2 - 20t - 10 = 0$$

from which

$$t = -0.45 \text{ sec} \quad \text{or} \quad +4.53 \text{ sec}.$$

The latter answer, happening *after* the stone was thrown, is obviously the one we want.

Most projectiles, however, are not considerate enough to move up and down in a vertical direction. What would have happened if the boy on the bridge had thrown the stone at 20 m/sec exactly horizontally, instead of vertically? If we neglect the frictional resistance of the air, the only force acting on the stone is the pull of gravity, and so its only acceleration is 9.8 m/sec^2 downward. This downward acceleration, since it has no horizontal component, cannot have any effect on the horizontal motion of the stone, which will continue at 20 m/sec until it strikes the water. Similarly, this horizontal component of the motion will have no effect on the gravitationally accelerated up-and-down component.

Figure 3-8 shows a ball being thrown exactly horizontally at the same instant that another ball is dropped. Since the initial velocity of the

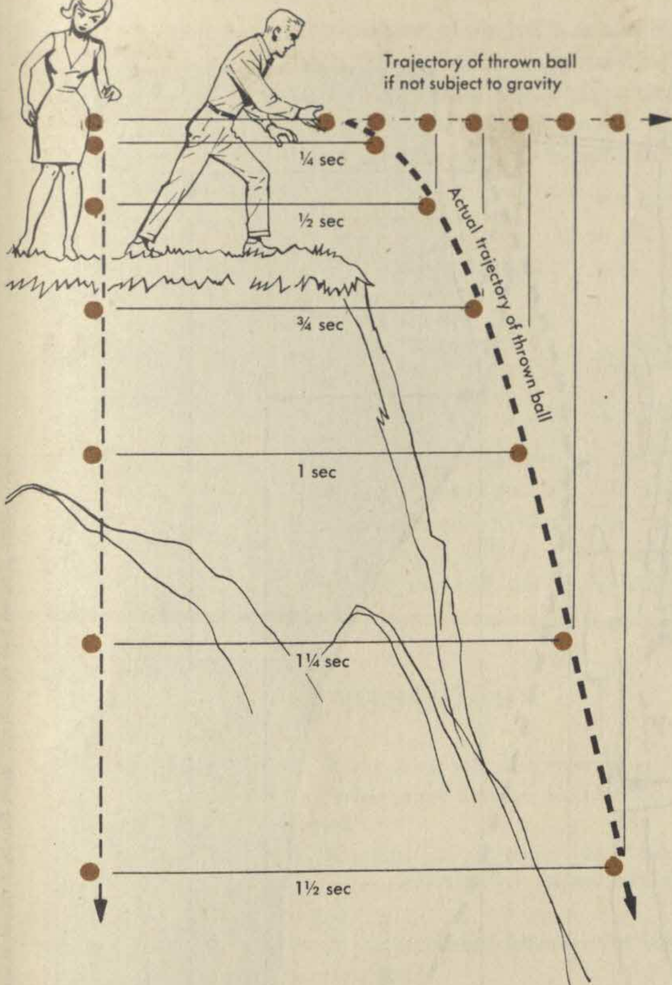


FIG. 3-8 Trajectory of a ball thrown horizontally.

thrown ball has no vertical component, its downward motion is the same as that of the dropped ball. The horizontal motion proceeds unchanged by the downward acceleration, so that the actual trajectory is the combination of the constant horizontal component and the accelerated downward component. Projectile problems can all be handled by considering the horizontal and vertical components of the motion separately and independently. As an example, Fig. 3-9 shows a golfer who has given a ball an initial velocity of 36 m/sec, 30° above the horizontal, from a tee that is 15 m below the surrounding terrain. How far will his drive carry? (This is the horizontal distance x on the drawing.) We can start by breaking v into its horizontal and vertical components: $v_h = 36 \cos 30° = 31.2$ m/sec, and $v_v = 36 \sin 30° = 18.0$ m/sec. The

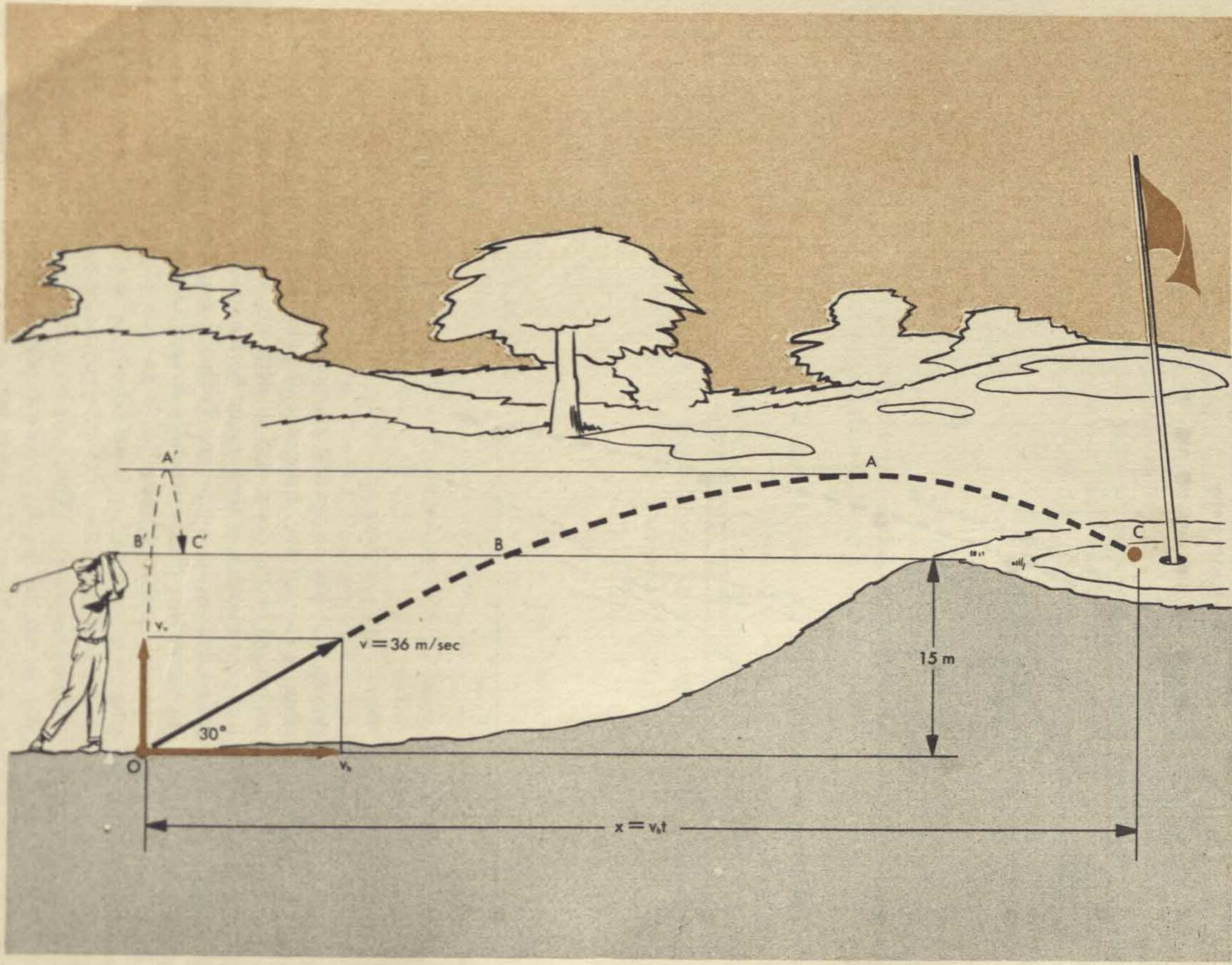


FIG. 3–9 Trajectory of a projectile launched with a velocity having both horizontal and vertical components.

vertical part of the ball's actual flight, $OBAC$, can be duplicated by the path $OB'A'C'$ of an imaginary ball thrown upward with a speed of 18 m/sec. How long will it take this imaginary ball to reach C'? When it does, the actual ball will be at C and will strike the ground. Dealing now with only the vertical component of motion we can write:

$$d = v_0t + \tfrac{1}{2}at^2$$
$$15 = 18t - 4.9t^2$$
$$4.9t^2 - 18t + 15 = 0$$

from which

$$t = \frac{18 \pm \sqrt{(18)^2 - 4 \times 4.9 \times 15}}{2 \times 4.9}$$
$$= 1.28 \text{ sec} \quad \text{or} \quad 2.40 \text{ sec}.$$

The answer 1.28 sec gives a time at which the ball is at the required level, 15 m above the tee, but it refers to point B, on the way up; the answer we need is 2.40 sec, the time at which the ball reaches C. During all this time, v_h will have been continuing on at 31.2 m/sec, and the ball will have traveled the horizontal distance

$$x = 2.40 \times 31.2 = 74.9 \text{ m}.$$

Questions

(3-1)

1. A car is capable of accelerating from rest to a speed of 60 mi/hr in 10 sec. (a) What is the average acceleration during these 10 sec? (b) How far does it travel?

2. A cyclist can accelerate from 5 mi/hr to 30 mi/hr in 40 sec. (a) What is his average acceleration? (b) How far does he travel during his 40 sec of acceleration?

3. A car accelerates from rest to 90 ft/sec in a distance of 400 ft. What is its acceleration?

4. An electron traveling 8×10^6 cm/sec enters an electric field which stops it in a distance of 4 cm. What is the acceleration of the electron?

5. An electron traveling 5×10^6 cm/sec passes through an electric field that gives it an acceleration of 10^{17} cm/sec^2. (a) How long a time will it take for the electron to double its initial speed? (b) Through what distance will the electron travel in this time?

6. In Question 5 you no doubt assumed the acceleration to be in the same direction as the velocity. Now work it again, assuming the direction of the acceleration to be opposite to the direction of the electron's initial velocity.

7. A train traveling 20 m/sec must reach a checkpoint 10 km distant in 6 min if it is to maintain its schedule. (a) What constant acceleration will bring the train to the checkpoint on time? (b) How fast will the train be traveling when it passes the checkpoint?

8. A beam of ions have a velocity of 2.0×10^6 cm/sec when they enter an accelerative electric field. It is necessary that they strike a cathode 30 cm away in just 8 μ sec. (a) What constant acceleration must the ions be given in order to accomplish this? (b) How fast will they be traveling when they strike the cathode?

9. A ball rolls down an inclined track 2.0 m long in 4 sec. (a) What is its acceleration? (b) What is its speed at the bottom of the track?

10. An electron, started from rest with uniform acceleration by an electric field, travels a distance of 2 cm in 2.5×10^{-8} sec before it strikes the anode. (a) What is its acceleration? (b) With what speed does it strike the anode?

(3-3)

11. If one were out to make money by buying pinto beans at one altitude and selling them for the same price at another altitude, should he buy or sell at the higher altitude location? (Weighing done on a spring balance.)

12. With gold selling everywhere at, say, \$600/lb, and presuming both buyers and sellers are willing to use your spring balance, should you buy it in Eagle City, Alaska, and sell it in the Canal Zone, or vice versa? How much would you have to buy and sell to make \$100 profit?

13. What net force is needed to give a mass of 450 gm an acceleration of 12 cm/sec^2? An acceleration of 12 m/sec^2?

14. What net force is needed to give a mass of 250 kg an acceleration of 30 cm/sec^2? An acceleration of 30 m/sec^2?

15. On a horizontal frictionless track, a force of 0.25 lb gives a body an acceleration of 4.0 ft/sec^2. What is the mass of the body?

16. A net force of 3.0 lb gives an acceleration of 15.0 ft/sec^2 to a body on a frictionless horizontal track. What is the mass of the body?

17. A constant force F pushes an originally motionless 300-gram mass a distance of 2 meters in 3 seconds, across a horizontal frictionless table. (a) What is its acceleration? (b) What is the force?

18. A constant force F pushes an originally motionless 12-pound mass a distance of 6 ft in 4 seconds, across a horizontal frictionless table. (a) What is its acceleration? (b) What is the force?

19. Work Question 18, if there is a coefficient of friction = 0.20 between the mass and the table.

20. Work Question 19, if there is a coefficient of friction = 0.25 between the mass and the table.

21. A 25-ton locomotive draws a 300-ton train at 45 mi/hr. (a) If the coefficient of friction between wheels and track is 0.20, what is the maximum acceleration the train can have? (b) Assuming that the cars, as well as the locomotive, have brakes, what distance will be required to bring the train to a stop?

22. A locomotive with a mass of 30 metric tons draws a 500 metric ton train at 60 km/hr. (a) If the coefficient of friction between wheels and track is 0.15, what is the maximum acceleration the train can have? (b) Assuming that the cars, as well as the locomotive, have brakes, what distance will be required to bring the train to a stop?

(3-4) **23.** How long will it take for a dropped stone to hit the bottom of a well 120 ft deep?

24. A coffeecup is dropped from a height of 4 ft. How long does it take it to hit the carpet?

25. A stone dropped from the top of a cliff hits the ground below 4 sec later. How high is the cliff?

26. A flowerpot dropped from the top of a building strikes the ground 3 sec later. How tall is the building?

27. On Planet X, a dropped stone falls 24 m in 2 sec. (a) What is g on Planet X? (b) What is the weight of a 5-kg mass on Planet X?

28. On Planet U, a dropped ball falls 56 ft in 3 sec. (a) What is g on Planet U? (b) What is the mass of an object that weighs 100 lb on Planet U?

29. An Atwood's machine is composed of masses of 1000 gm and 1020 gm suspended from a frictionless pulley. (a) What will be the downward acceleration of the heavier mass? (b) When released from rest, how long will it take the heavier mass to descend 200 cm?

30. An Atwood's machine has weights of 20 lb and 18 lb on a cord passing over a frictionless pulley. (a) What will be the downward acceleration of the 20-lb weight? (b) When released from rest, how long will it take the heavier weight to descend 4 ft?

31. What is the tension in the cord in Question 29?

32. What is the tension in the cord in Question 30?

33. Consider a horizontal frictionless table with a frictionless pulley mounted at its edge. From a 2000-gm weight on the table, a horizontal string passes over the pulley; the other end of the string hangs downward from the pulley, and attached to it is a 250-gm weight. What is the downward acceleration of the hanging 250-gm weight?

34. Consider a horizontal frictionless table with a frictionless pulley mounted at its edge. From a 2.50-kg weight on the table, a horizontal string passes over the pulley; the other end of the string hangs downward, and attached to it is a mass of 0.60 kg. What is the acceleration of the weight on the table?

35. If the coefficient of friction between weight and table in Question 33 were 0.10, how fast would the 2000-gm weight be moving at the end of 1 sec, starting from rest?

36. If the coefficient of friction between weight and table in Question 34 were 0.15, how far would the hanging weight descend in 2 sec?

37. What is the tension in the string in Question 35?

38. What is the tension in the string in Question 36?

39. A 180-lb man stands on a spring scale in an elevator moving downward. If the elevator decelerates 4 ft/sec^2, what will be the reading on the scale?

40. A 50-kg man stands on a spring scale in an upward-moving elevator. The elevator decelerates 2 m/sec^2 as the elevator stops. What is the scale reading during the deceleration? (This scale is calibrated in newtons; scales calibrated in these units are to be found only in physics texts and physics laboratories.)

41. A man stands on a spring scale in a moving elevator in Paris. As the elevator comes to a stop the scale reading changes from 60 kg to 45 kg, then back to 60 kg when the elevator stops moving. (a) Was the elevator going upward or downward? (b) What was its acceleration while stopping?

42. A man stands on a spring scale in a moving elevator. As the elevator stops, the scale reading goes from 160 lb to 190 lb, then back to 160 lb when the elevator stops moving. (a) Was the elevator going upward or downward? (b) What was its acceleration while stopping?

(3-5)

43. An object slides down a frictionless inclined plane making an angle 20° with the horizontal. What is the acceleration of the object along the plane?

44. A frictionless plane is inclined at an angle of 25° with the horizontal. What is the acceleration of an object sliding down this plane?

45. In Question 43, include a coefficient of friction = 0.20.

46. In Question 44, include a coefficient of friction = 0.15.

47. A 1000-lb boat is slid down a ramp making an angle of 15° with the horizontal; the coefficient of friction is 0.30. Will the boat slide into the water of its own accord? If not, what push (parallel to the ramp) will be required?

48. A 200-lb package marked "FRAGILE" slides down a 20° loading ramp; the coefficient of friction is 0.20. With what force, parallel to the ramp, must the loaders restrain the package so it will slide down at a constant speed?

49. A book will slide at constant speed down an incline that makes an angle θ with the horizontal. Show that the coefficient of friction between book and incline is equal to $\tan\theta$.

50. The coefficient of friction between a block and an inclined plane is μ. If the block slides down at a constant speed, what is the tangent of the angle the plane makes with the horizontal?

51. The upper end of a ramp 13 ft long is 5 ft higher than its low end. A man flicks a package of cigarettes up the ramp ($\mu = 0.15$) with such a speed that it comes to rest exactly at the top of the ramp. What was the initial speed of the package?

52. A bale of rubber latex will not slide down a ramp inclined 15° with the horizontal, because the coefficient of friction is 0.40. What velocity must the bale be given at the top of the ramp to cause it to slide down and come to rest exactly at the bottom? The ramp is 20 ft long.

(3-6)

53. A man stands by a well 120 ft deep and throws a stone upward at 20 ft/sec. The stone comes down into the well. How long a time after it is thrown will the stone strike the bottom of the well?

54. A man atop a building 50 m high shoots a stone upward with a sling-shot at a speed of 30 m/sec. How long before the stone lands on the ground at the base of the building?

55. On Planet X, a man throws a 500-gm mass upward with a speed of 20 m/sec, and catches it as it comes down 20 sec later. What does the 500-gm mass weigh?

56. On Planet Y, a mass of 2 kg is thrown upward with a speed of 10 m/sec, and is caught as it comes down 8 sec later. What is the weight of the 2-kg mass?

57. A baseball is thrown with a speed of 120 ft/sec in a direction 37° above the horizontal. (a) To what height will the ball rise? (b) How far (measured along the horizontal) will the ball travel?

58. A golf ball is driven with a speed of 40 m/sec in a direction 30° above the horizontal. (a) To what height will the ball rise? (b) What horizontal distance will the ball travel?

59. A football is punted at an angle of 60° above the horizontal. (a) With what speed did it leave the punter's toe if it travels 50 yards, measured horizontally on the field? (b) To what maximum height does the ball rise?

60. A soccer punt starts at an angle of 45° above the horizontal, and travels a horizontal distance of 30 yards. (a) To what maximum height does the ball rise? (b) With what speed did it leave the player's toe?

61. A plane, diving at an angle of 30° below the horizontal with a speed of 150 m/sec, releases a bomb when it is 1000 m above the ground. (a) How many seconds pass before the bomb strikes the ground? (b) How far beyond the point of release will the bomb strike the ground?

62. A plane is climbing at an angle of 30° above the horizontal, at a speed of 300 m/sec. When it is at an altitude of 1200 m it releases a bomb. (a) How many sec pass before the bomb strikes the ground? (b) How far beyond the point of release does the bomb strike?

chapter four

Energy and Momentum

4-1 Work and Potential Energy

Let us return now to two cases discussed earlier: (1) two forces applied to the ends of a solid bar supported at a certain point in between, and (2) two forces applied to the pistons of different diameters in two connected cylinders. We assume here that the bar and the pistons have a negligibly small weight and that friction can be disregarded. These two cases of lever arms or areas with ratios of 1 : 2 are shown in Fig. 4-1A and B.

From our knowledge of levers and Pascal's law, we know that a downward force on the left side of either of these machines will be able to raise a weight on the right side that is twice as great as the force. This "magnification" of a force—the ratio of the force the machine can exert to the force which must be exerted on the machine—is called the *mechanical advantage* of the machine; in each of the above cases the mechanical advantage is 2.

Along with this advantage, however, is a disadvantage. For a downward push, of, say, 1 ft on the left side, the weight on the right is raised

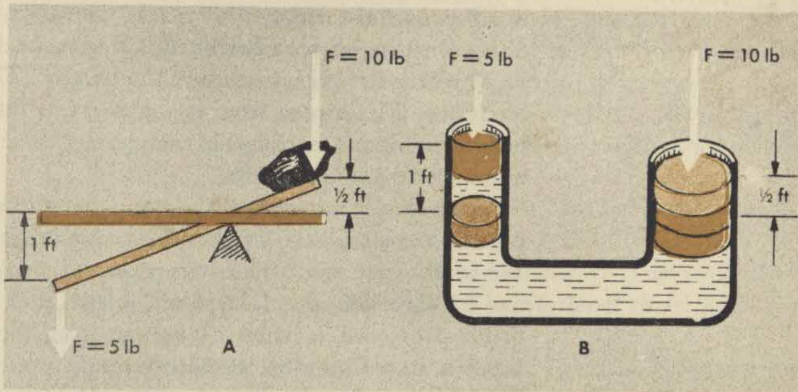


FIG. 4–1 Equilibrium conditions for the forces applied to the ends of a lever, to the pistons of a hydraulic press, or to any other machine of equal mechanical advantage.

in both cases only $\frac{1}{2}$ ft, so what we have gained in force we have lost in distance. In other words, the product of force and distance is the same on both sides. This product of force times the distance moved *in the direction of the force* is the physicist's definition of *work*. It must, of course, be expressed in units that have the dimensions of a force times a distance, such as a foot-pound, or a dyne-centimeter, or an ounce-inch. In the simple examples of Fig. 4-1, we might consider a force of 5 lb on the left side moving through a distance of 1 ft, and thus doing 5 ft-lb of work on the machines. The machines in turn exert a force of 10 lb to lift a weight on the right side through a distance of $\frac{1}{2}$ ft, thereby doing $10 \times \frac{1}{2} = 5$ ft-lb of work on the stone or other restraining force.

This may also be taken as an example of the very important principle of *conservation of energy*. This principle says that work, or *energy* (which is simply the ability to do work), never appears from nowhere and never vanishes. The 5 ft-lb of work done on the machine merely transformed into the equal amount of work the machine did in raising the stone. This cannot be the end of the story, though, for what has happened to the work done on the stone?

The stone was raised, against the pull of gravity, $\frac{1}{2}$ ft higher than it was before, and it is now capable of exerting a force of 10 lb (its weight) through a distance of $\frac{1}{2}$ ft, in returning to its original position. As long as the stone is elevated, the 5 ft-lb of work done on it is stored, ready to be released whenever the stone is lowered. It has, in other words, the potentiality of doing this amount of work, which is called its *gravitational potential energy*.

The work an elevated object can do is of course (like all work in the physics sense) measured by force times distance; in the case of gravita-

tional potential energy, the force is the weight of the object, and the distance is the object's height. But this immediately raises a question. From where should we measure the height? The answer is a very generous one. Measure it from any place you find convenient! We shall be interested only in *changes* in energy, and these changes will be the same no matter what we choose.

In the simple example of the stone and the lever, it might have been easiest to measure the height h from the original position of the stone. We would then say it had zero potential energy (PE) in this location, and 5 ft-lb of PE after it had been raised, thus giving it a 5-ft-lb *increase* in PE. However, a homesick student from the seashore solving such a problem in a Colorado laboratory might prefer to refer his zero to sea level. If, for example, the elevation of the stone was originally 5612.3 ft above sea level, the student would be entitled to say that the stone had a PE of $5612.3 \times 10 = 56{,}123$ ft-lb. The work of the lever would raise it to an elevation of 5612.8 ft and a PE of 56,128 ft-lb, again giving the same answer—that its PE had increased by 5 ft-lb.

Another student might prefer to use the ceiling of the laboratory as his zero. He would run into only a slight complication, because if the PE were zero at the ceiling, it would be negative on the laboratory bench. If the stone were originally 9 ft below the ceiling, its PE would be $-9 \times 10 = -90$ ft-lb; when raised, it would be -85 ft-lb. Since -85 is greater than -90, he also would conclude that the gravitational potential energy had been increased by 5 ft-lb.

The notion of potential energy is not necessarily associated only with the force of terrestrial gravity. A tightly wound spring or a gas compressed in a metal cylinder is also able to produce mechanical work that can be measured in the same units. Potential energy in chemical, rather than mechanical, form is stored in an automobile fuel tank filled with gasoline or in a charge of high explosive in an artillery shell. Potential energy in still another form lies in the nuclear energy of the plutonium core of an A-bomb.

4-2 Kinetic Energy

We know that if a massive body is moving rapidly, it can do a certain amount of work (or damage) that it could not do when standing still. *This ability of moving bodies to produce mechanical work because they are moving is known as their kinetic energy* and can be expressed either in terms of the work done to bring them into that state of motion or in terms of the work they can do before coming to rest. We can investigate the situation by means of the simple experiment shown in Fig. 4-2. A light carriage that can be loaded with a variable amount of heavy weights is rolled with a certain velocity along a table. A string attached to the carriage passes over a pulley and its other end is attached to a weight resting on the

floor at the side of the table (Fig. 4-2A). When the string tightens, the weight will be lifted from the floor and will reach a certain maximum elevation at the moment the carriage comes to rest (Fig. 4-2B). If we double the load on the carriage, keeping its initial velocity the same, the weight will be raised twice as high (Fig. 4-2A). If, on the other hand, we use the same load but propel the carriage with twice the velocity, the weight will be raised to four times the height recorded in the first experiment. Similarly, if we triple the load of the carriage, the weight will be raised to a triple height; and if we triple the velocity, the weight will be lifted nine times as high. Thus we must conclude that ***the kinetic energy of a moving object is proportional to its mass and to the square of its velocity.***

By making use of the conservation of energy principle, we can justify these experimental results by a bit of straightforward pencil-and-paper work. Suppose we again take a mass of m gm on a frictionless cart (Fig. 4-3). If we accelerate the cart by pulling on it with a constant force of F dynes and continue to apply this force until it has moved a distance d, we shall have done Fd dyne-cm of work on the cart. According to the conservation of energy principle, this work must be accounted for in some way; the mass has not been raised, so its gravitational potential energy is not increased, and none of the work has been lost in overcoming friction. We must conclude that the kinetic energy (KE) of the mass is just equal to the work Fd:

$$KE = Fd.$$

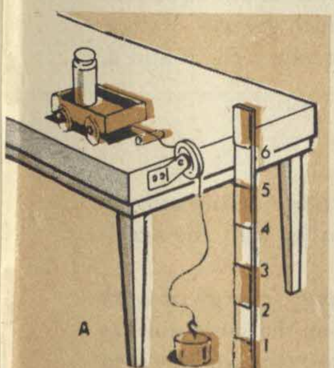
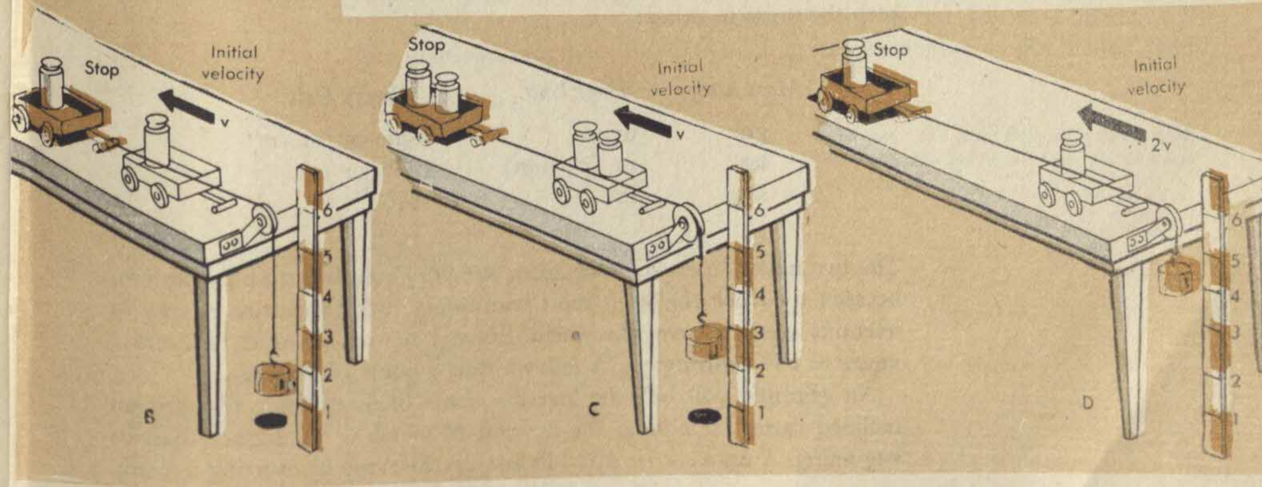


FIG. 4-2 The work that can be done by a moving object is in proportion to its mass, and to the square of its velocity.



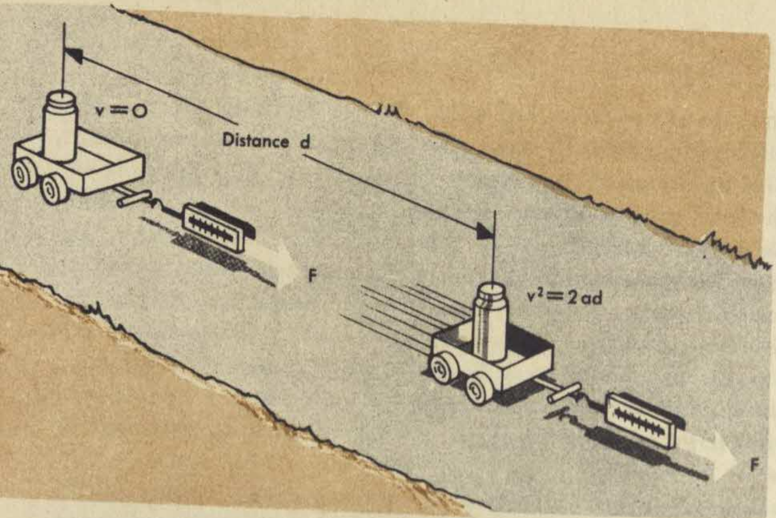


FIG. 4–3 Conversion of work into kinetic energy.

For bodies starting from rest, where $v_0 = 0$, we know from our equations of uniformly accelerated motion that $v^2 = 2ad$, which gives $d = v^2/2a$. We also know that $F = ma$, and substituting these values for F and d, we get

$$\mathrm{KE} = Fd = ma \times \frac{v^2}{2a} = \frac{1}{2}mv^2.$$

This important general expression for the kinetic energy of any moving body confirms the experimental findings made earlier with the accelerated carts.

The "force" part of the units used for kinetic energy (energy or work equals a force times a distance) must be expressed in the same units as in $F = ma$, which is to be expected, since Newton's second law was used in deriving the expression for KE. Here is a tabulation that may help keep the units in mind:

Mass Unit	*Force Unit*	*Energy Unit*
gm	dyne	dyne-cm or *erg*
kg	nt (newton)	nt-m or *joule*
slug	lb	ft-lb

The two new names, *ergs* and *joules*, are very convenient abbreviations, because they are the units most commonly used to express energy in scientific work all over the world. Since 1 newton $= 10^5$ dynes, and 1 meter $= 10^2$ centimeters, it follows that 1 joule $= 10^7$ ergs.

An example will help tie together some of these ideas. Imagine an inclined ramp 50 ft long, the far end of which is 6 ft higher than its beginning. A man shoves a 240-lb box up the ramp by exerting a steady

push of 80 lb. Let us assume that the force of friction between the box and the ramp is 40 lb. How fast will the box be moving when it reaches the top of the ramp? The total amount of work the man does can be determined by multiplying the force of his push by the distance through which he pushes the box:

$$Fd = 80 \times 50 = 4000 \text{ ft-lb.}$$

We must now consider what becomes of this 4000 ft-lb of work. Since the force of friction is given as 40 lb, the work that goes into overcoming friction and that is thus converted into heat is $40 \times 50 = 2000$ ft-lb. When it reaches the end of the ramp, the box will have been raised through a vertical distance of 6 ft, and consequently it will have gained $240 \times 6 = 1440$ ft-lb of potential energy.

This will account for $2000 + 1440 = 3440$ ft-lb of the 4000 ft-lb of work the man has done. Left over, we have $4000 - 3440 = 560$ ft-lb of work which can only have gone into increasing the speed and thus the kinetic energy of the box. We need now only determine how fast the box must be moving to give it 560 ft-lb of kinetic energy. To do this, we can compute directly from $KE = \frac{1}{2}mv^2$ (remembering that since the 560 is in foot-pounds, the mass must be expressed in slugs):

$$560 = \frac{1}{2} \times \frac{240}{32} \times v^2$$

$$v^2 = \frac{560 \times 2 \times 32}{240} = 149.3$$

$$v = 12.2 \text{ ft/sec.}$$

Problems concerning work may often involve a force which acts in a direction different from the direction of the motion of the body it is acting on. In Fig. 4-4A is a cart being pulled up a hill by a force F, which is not parallel to the direction of motion. Figure 4-4B shows F resolved into two components: F_{perp}, perpendicular to the motion, and

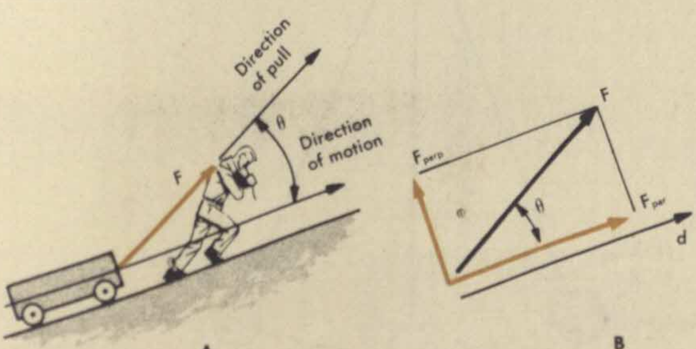


FIG. 4–4 The work done by a force acting in a direction different from the direction of motion.

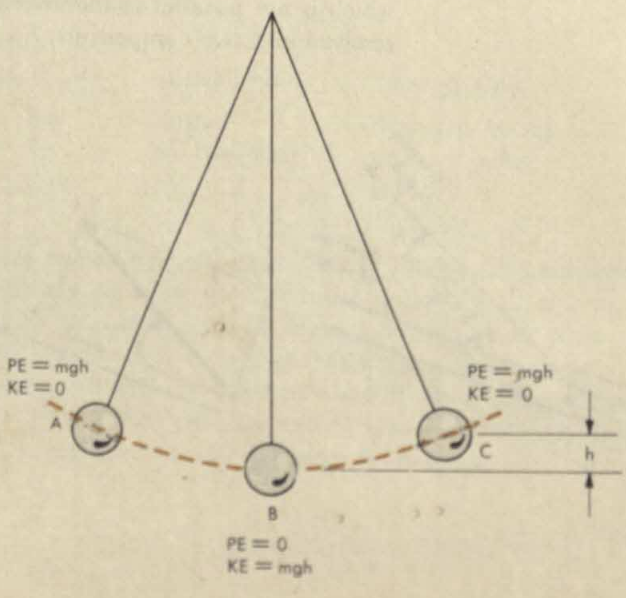
F_{par}, parallel to the motion. The work done is F_{par} times the distance d. Since F_{par} is $F \cos \theta$, we see that the general expression for the work done by a force is $Fd \cos \theta$. Because F_{perp} has no component in the direction of d, it does not do any work at all.

4-3 Energy Interchanges

In many examples of mechanical motion, there is an interchange between kinetic and potential energies. Thus, if we hold a Ping Pong ball in our hand some distance above the floor, it has potential energy but no kinetic energy. If we release it, it falls faster and faster toward the floor, and when it reaches the floor, it will have no potential energy left; all of it will have been turned into kinetic energy. At the moment of impact with the floor, the ball will stop for a split second, and all its kinetic energy will be turned into the potential energy of the elastic deformations in its body. (In the case of an inelastic lead ball, the kinetic energy will be transformed into heat—by a process that we shall explain later—and it will not bounce.) The elastic energy is then changed back into kinetic energy, and this sends the Ping Pong ball up into the air, with the result that the kinetic energy is turned into gravitational potential energy. The ball will rise to approximately its original height. This process is repeated again and again until the friction forces gradually rob the system of its initial energy and the ball comes to a standstill on the floor.

A pendulum is another example of interchange of energy between PE and KE. Consider the pendulum shown in Fig. 4-5. At the ends of

FIG. 4-5 Conversion of potential energy into kinetic energy, and kinetic energy into potential energy in a swinging pendulum.



its swing (A and C), the pendulum bob is momentarily motionless and therefore has no kinetic energy. However, at these two points it is at its maximum elevation, a distance h above its elevation at B, in the center of its swing. If we take B as our zero, then at A and C its gravitational potential energy is mgh, while its kinetic energy is zero. At B the situation is reversed; its PE is zero, and its speed and KE are maximum. We can easily determine the speed of the pendulum bob at B by calling on the principle of the conservation of energy. If we neglect the small loss of energy due to air friction and the bending of the supporting string, we can conclude that the gain in KE from A to B must exactly equal the loss in PE:

$$\tfrac{1}{2}mv^2 = mgh$$

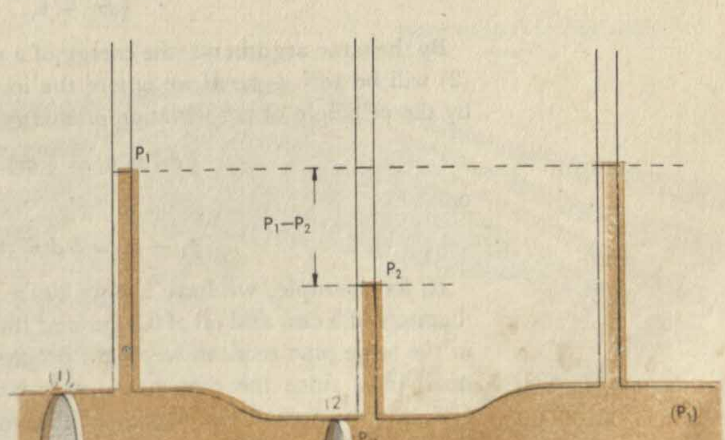
or

$$v = \sqrt{2gh}.$$

4-4 Bernoulli's Principle

Consider a liquid flowing through a pipe of varying size (Fig. 4-6). Since liquids are almost incompressible, the volume per second passing through the large section at (1) must be the same as that passing through the small section at (2). In order to do this, the liquid must speed up as it enters the small section; that is, between (1) and (2) the liquid is accelerated to the right. In order to provide the force to cause this acceleration, the pressure at (1) must be greater than the pressure at (2). Similarly, the liquid must be decelerated when it leaves (2) and

FIG. 4-6 The pressure is lower in the narrow part of the tube, where the fluid moves faster.



reenters the large section of pipe to the right; this means there must be a retarding force toward the left, or that, again, the pressure is larger in the larger pipe. This fact can be easily demonstrated by attaching narrow vertical pipes to the three parts of our horizontal pipe, as shown in Fig. 4-6. The water in the middle pipe will stand lower and thus indicate a lower pressure. The statement that ***in the regions where the velocity of fluid is smaller, the pressure is higher, and vice-versa,*** is known as the *principle of Bernoulli*, after a Swiss physicist, Daniel Bernoulli (1700–1782), who discovered it.

The principle of conservation of energy enables us to deal quantitatively with Bernoulli's principle. Since the same amount of liquid per second must flow through every part of the pipe, we can see that

$$v_1A_1 = v_2A_2.$$

Imagine at location (1) a unit volume of the liquid (1 cm^3, or 1 ft^3, or 1 m^3). Its mass is numerically equal to its density d, and it has a kinetic energy of $\frac{1}{2}dv^2$. Our pipe is horizontal, so the gravitational potential energy of the liquid will not change as it moves along the pipe. However, there is another form of energy, due to pressure, that we must take into account. Imagine a hole 1 cm square cut in the side of the pipe. Now into this hole stuff a 1-cm cube of the liquid. (This is *very* difficult to do and you will be much better off just imagining it.) The force against the 1-cm^2 face of the cube is numerically equal to the pressure, and you must push it in a distance of 1 cm, which makes the work you do also numerically equal to the pressure p. This amount of work p can obviously be reclaimed by allowing the cube to emerge into the open again, and thus p represents the energy the 1 cm^3 of fluid has because of its pressure. The total energy (ignoring gravitational PE) of the unit volume is then

$$\tfrac{1}{2}dv_1^2 + p_1.$$

By the same arguments, the energy of a unit volume of liquid at point (2) will be $\frac{1}{2}dv_2^2 + p_2$. If we ignore the loss of energy by friction, then by the principle of conservation of energy we get

$$\tfrac{1}{2}\,dv_1^2 + p_1 = \tfrac{1}{2}\,dv_2^2 + p_2$$

or

$$p_1 - p_2 = \tfrac{1}{2}\,d(v_2^2 - v_1^2).$$

If, for example, we have a pipe 10 cm in diameter narrowing to a diameter of 5 cm, and oil of 0.8 gm/cm^3 flowing at a speed of 100 cm/sec in the large pipe section, we might proceed somewhat like this: In the small pipe, since the diameter is only half as great, the area will be $(\frac{1}{2})^2$, or $\frac{1}{4}$ the area of the large pipe. This means that the velocity v_2 must be 4 times v_1, or 400 cm/sec. On the right side of the equation, d is given in gm/cm^3, and the speeds in cm/sec, which is consistent with the

left side, in which the pressures are in dynes/cm². Numerically,

$$p_1 - p_2 = \tfrac{1}{2} \times 0.8 \times (400^2 - 100^2) = 6 \times 10^4 \text{ dynes/cm}^2.$$

This scheme is extensively used to measure the flow of fluids. A gauge reads the pressure difference between the normal pipe and a special narrowed section; if we know the ratio of pipe diameters, we need merely a reversal of the example given above to find the flow velocity.

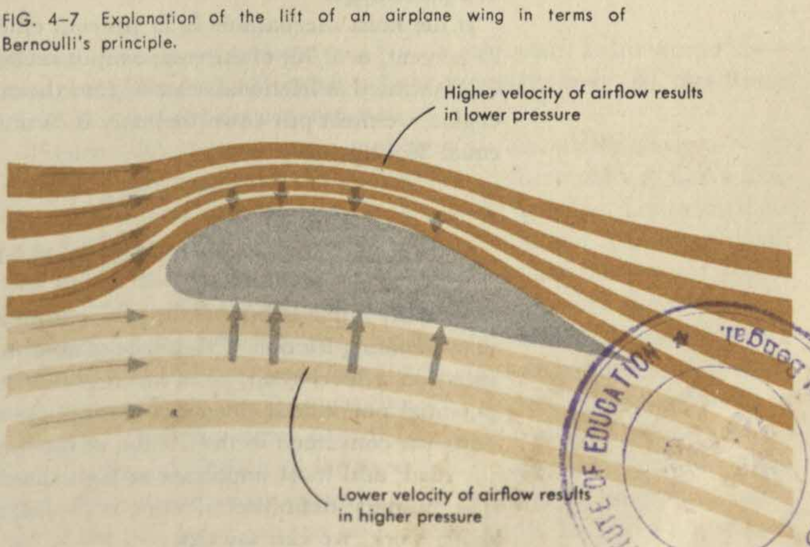
Bernoulli's principle is quite general and applies to all kinds of fluid motion. Consider, for example, the stream of air around the wing of a flying plane. The profile of the wing and the lines of air flowing around it are shown in Fig. 4-7. Airplane wings are shaped in such a way that the total distance traveled by the air flowing over the wing is longer than that of the air flowing under it. Thus the velocity of airflow above the wing must be higher, and the pressure correspondingly lower, than that of the airflow under the wing. This difference of pressure above and below the wing results in an upward force that helps to support the flying airplane in the air.

4-5 Power

Webster's New Collegiate Dictionary gives ten separate meanings for the word "power." Although the nine other meanings are perfectly legitimate in ordinary conversation, in physics we must confine ourselves to only one. Power is *the* ***rate*** *of doing work or the* ***rate*** *at which energy is converted from one form into another.*

As an example, we can figure the power needed to raise a 5000-lb

FIG. 4-7 Explanation of the lift of an airplane wing in terms of Bernoulli's principle.



elevator a distance of 120 ft in 30 sec. The total work that must be done is 5000 lb $\times$ 120 ft = 600,000 ft-lb. Even a small motor could do this amount of work if it were allowed a long enough time. Our problem, however, says that the job must be done in 30 sec, so the hoist must work at the rate of

$$\frac{600{,}000}{30} = 20{,}000 \text{ ft-lb/sec.}$$

Ordinarily, for such things as hoists and engines, a larger unit of power is used: 1 *horsepower (hp) equals 550 foot-pounds per second (ft-lb/sec)*. So, if the hoist mechanism were 100 percent efficient—that is, if none of its work were lost in overcoming friction—the elevator would require an engine of

$$\frac{20{,}000}{550} = 36.4 \text{ hp.}$$

No perfectly frictionless machine exists, however, and some allowance must be made for the work lost by being converted into heat by friction. In any machine, the work it does (its *output*) will be less than the energy supplied to it, or the work done on it (its *input*), because of these frictional losses. The term *efficiency*, when applied to a machine of any kind, means merely the ratio of its output to its input:

$$\text{efficiency} = \frac{\text{output}}{\text{input}}.$$

Efficiency must obviously always be less than 1, and is commonly given as a percentage.

If the hoist mechanism is 75 percent efficient, this means that only 75 percent, or 0.75, of the engine input is usable in raising the elevator; 0.25 is wasted as frictional heat. To find the actual power of the required engine, we must put down that only 0.75 of the input to the hoist must equal 36.4 hp, or

$$0.75x = 36.4$$

$$x = \frac{36.4}{0.75} = 48.5 \text{ hp.}$$

In many applications, *all* the power supplied to a machine is used up in overcoming friction. This is true of an automobile moving at uniform speed on a level road, or of an airplane in horizontal flight. Neither potential nor kinetic energy is being increased, and the entire engine output is consumed in the friction of moving parts, of the tires against the road, and most important at high speed, in the friction of the air. The primary definition of work is $F \times d$; since power is the rate of doing work, we can say that:

$$\text{power} = \frac{\text{work}}{\text{time}} = \frac{Fd}{t} = F \times \frac{d}{t} = Fv.$$

Therefore, if an automobile engine develops 200 hp when the car is moving at 75 mi/hr, the total force of friction can be easily figured. To get everything expressed in a consistent set of units, we write

$$200 \text{ hp} = 200 \times 550 = 110{,}000 \text{ ft-lb/sec}$$

and

$$75 \text{ mi/hr} = \frac{75 \times 88}{60} = 110 \text{ ft/sec}.$$

Substitution of these figures into $P = Fv$ gives us

$$110{,}000 \text{ ft-lb/sec} = F \text{ lb} \times 110 \text{ ft/sec}$$

or

$$F = 1000 \text{ lb force of friction.}$$

In the widely used MKS system of units, power is, of course, expressed in joules per second. This unit of power has been given the name *watt* in honor of James Watt, the eighteenth-century British inventor of the steam engine. *One watt equals a power of one joule per second.* Although the watt is frequently used to describe electrical power, its definition is fundamentally mechanical and, like any other power unit, can be used to measure the rate of conversion of any kind of energy. We might, for example, determine the power expended by a 70-kg man who takes 12 sec to run up a flight of stairs 6 m high. His weight is $70 \times 9.8 = 686$ newtons. The total work done is 686 nt $\times$ 6 m = 4120 nt-m, or 4120 joules. The power exerted, then, is 4120 joules/12 sec = 342 joules/sec, or 342 watts.

4-6 Action and Reaction

Newton's third law of mechanics—at least as important as his second law—says that ***if one body (A) exerts a force on another body (B), then B must exert an equal and opposite force on A.***

Figure 4-8A shows a nurse pushing a perambulator along a rough walk at a constant speed. All goes well, and three pairs of forces illustrating Newton's third law can be seen on the drawing. Let us say there is a 7-lb frictional force between the perambulator and the ground; this means the wheels must push *forward* against the ground with a force of 7 lb, while the ground pushes *backward* against the wheels with an equal 7-lb force. Similarly, the push of the nurse on the perambulator and the push of the perambulator back on her hands are another equal and opposite pair, each amounting to 7 lb, and exactly opposite in direction. There is another pair of 7-lb forces between the nurse's shoes and the ground. In fact, so far as we know, there is no such thing as a single force; forces *always* occur in equal and opposite pairs, each member of the pair being exerted on a *different* body. In Fig. 4-8A, we can see two equal and opposite forces acting on the nurse; one exerted by the perambulator, and one exerted by the ground. These are *not* a third law pair—they are both acting on the same body; each one is a member

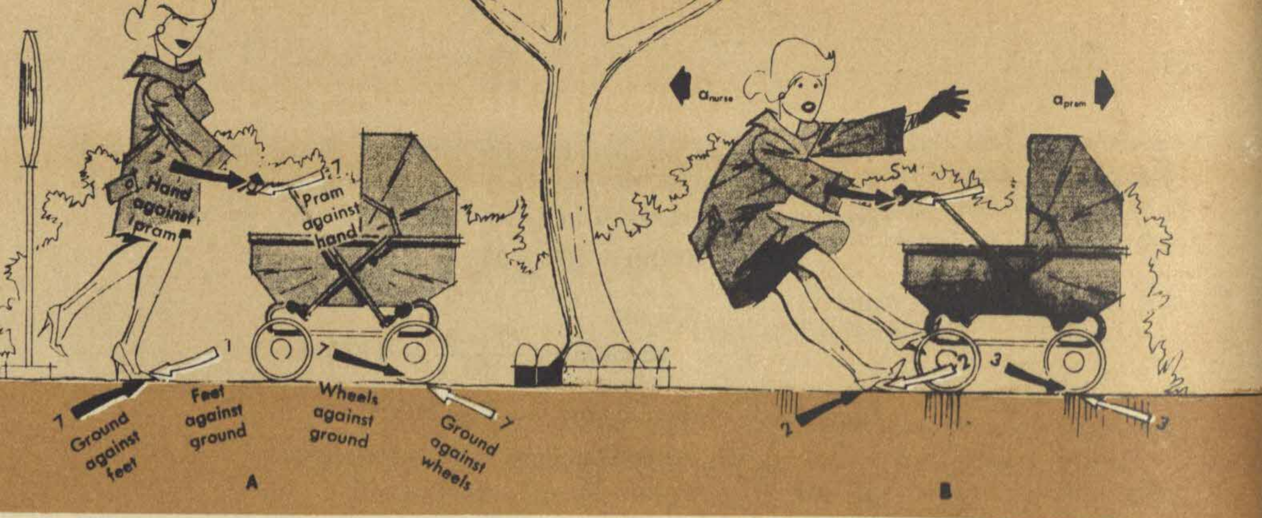


FIG. 4–8 Newton's third law: action and reaction are equal and opposite.

of a different pair. In this example, they happen to be equal and opposite only because the nurse happens to be pushing the perambulator at a constant speed on a level path.

In Fig. 4-8B, the situation is different. The rough gravel has given way to a patch of ice. The nurse, not noticing, has continued her 7-lb push; but on the slick surface the maximum forces of friction have been reduced to 3 lb for the perambulator, and to 2 lb for the smooth soles of the nurse's shoes. Again the frictional forces and the hand-handle forces occur in equal and opposite pairs, but the separate forces on the nurse and on the perambulator are no longer balanced. The nurse now experiences a net force of $7 - 2 = 5$ lb to accelerate her backward; the $7 - 3 = 4$-lb net force on the perambulator sends it shooting forward.

Since the forces of action and reaction between two objects are equal to one another and since the acceleration communicated to any object by a given force is inversely proportional to the mass of that object, the acceleration of the more massive of the two interacting objects will be smaller than that of the lighter one. When we shoot a rifle, for instance, the powder gases in the barrel force the bullet out toward the target, but they also push equally strongly toward the rear of the barrel and produce what is known as *recoil*. The recoil in an ordinary rifle is relatively small because the rifle is much more massive than the bullet, but even then, it is quite appreciable.

4-7 Momentum

We have already looked at one very useful quantity that depends on mass and motion—kinetic energy. Another quantity that is at least

equally important is *momentum*, which is merely the product of the mass of a body times its velocity:

$$\text{momentum} = mv.$$

As we shall see when we look at the concepts of Einstein's Special Theory of Relativity, the mass of a body increases as its velocity increases, but this change is negligibly small except at really enormous speeds. For the present, let us confine ourselves to modest speeds of no more than a few thousand miles per second, so that the mass of a body can be considered to be constant. Under these conditions, a change in the momentum of a body must mean a change in its velocity, and vice versa.

Now go back to Newton's second law, although its connection with momentum may not be apparent at first glance:

$$F = ma.$$

Acceleration is the rate of change of velocity, so we can write, instead of a, $(v_t - v_o)/t$. Substituting this value for a in Newton's second law, we get

$$F = \frac{m(v_t - v_o)}{t} = \frac{mv_t - mv_o}{t}.$$

Put into words, this statement that the force applied to a body equals the rate of change of its momentum is actually the way Newton himself stated his second law. If we multiply this equation through by t, we get

$$Ft = mv_t - mv_o.$$

On the left side of the equation, the quantity Ft—the product of a force and the time during which it acts—is called *impulse*; on the right side of the equation is the change in the body's momentum. The fact that the impulse given a body equals its change of momentum leads us directly to the very important concept of the *conservation of momentum.*

Consider two billiard balls colliding on a smooth table. During the fraction of a second the balls are in actual contact, each is slightly compressed, and in springing back to spherical shape each exerts a force on the other that will change its speed and direction. From Newton's third law, we know that these forces are at all times exactly equal and exactly opposite in direction, and the time during which these collision forces act is obviously identical for each ball. Thus each ball, during the collision, receives an impulse equal to the impulse received by the other ball, but opposite in direction. (Impulse and momentum are both vector quantities, since each is made up of a vector multiplied by a scalar.) It follows, then, that each ball undergoes a change in momentum equal and opposite to the change of momentum of the other ball. So if these changes are added (vectorially, remember!), they add up to zero.

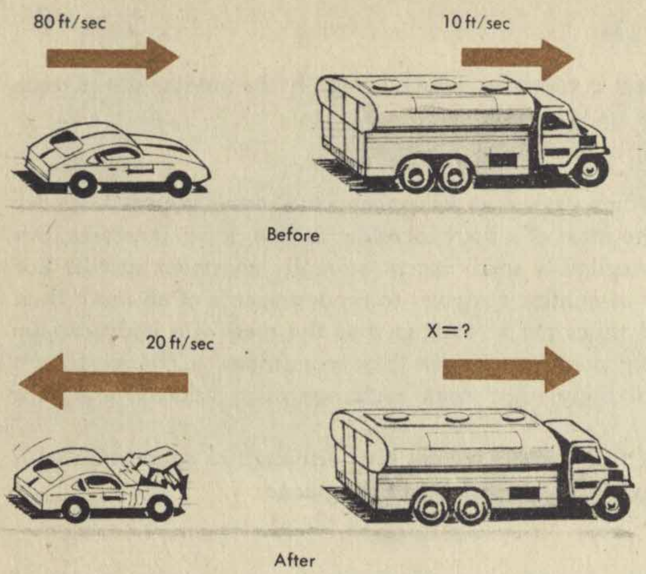


FIG. 4-9 Conservation of momentum in a collision.

The fact that the total momentum change is always zero allows us to make the statement that ***in any collision or other interaction between bodies, the total momentum of the interacting bodies (considered vectorially) is the same afterward as it was before,*** which is one way of stating the law of *conservation of momentum.*

As an example, consider a 2000-lb car traveling at 80 ft/sec that crashes into the rear of a 16,000-lb truck moving in the same direction at 10 ft/sec. The small car bounces backward from the collision with a speed of 20 ft/sec. What is the speed of the truck after the impact? (See Fig. 4-9.) In this example, the motions are in one dimension along a straight line, and we can take care of vector directions by assigning a + sign for motion to the right and a − sign for motion to the left. Before the collision, the total momentum of the two vehicles is

$$80 \times 2000 + 10 \times 16{,}000 = 320{,}000 \text{ lb-ft/sec}.$$

Afterward, it is $-20 \times 2000 + 16{,}000x$. If we set the momentum before the collision equal to the momentum after, we get

$$320{,}000 = 16{,}000x - 40{,}000$$

or

$$x = 22.5 \text{ ft/sec}.$$

If you were to check the total kinetic energy of the two cars before and after the collision, you would find that more than 38 percent of the original KE had been lost; this lost KE was used up (and converted into heat) in the work devoted to crumpling fenders, etc. When some mechanical energy is lost in this way, the collision is called *inelastic.* There is no such thing as a perfectly elastic collision between ordinary bodies—that is, a collision in which no energy is lost. Glass or hard steel balls may come fairly close to it, however. A golf ball dropped on the

pavement will bounce for a number of times before coming to rest, but each bounce will be to a lesser height than the previous one, thus giving us an indication that the ball has lost energy in each collision with the earth. At each impact with the pavement, the ball is slightly flattened. This causes the fibers in the ball to rub against each other and thus to use up energy by converting it to frictional heat which cannot be recovered. (We should add, however, that collisions between individual atoms and molecules are generally perfectly elastic.)

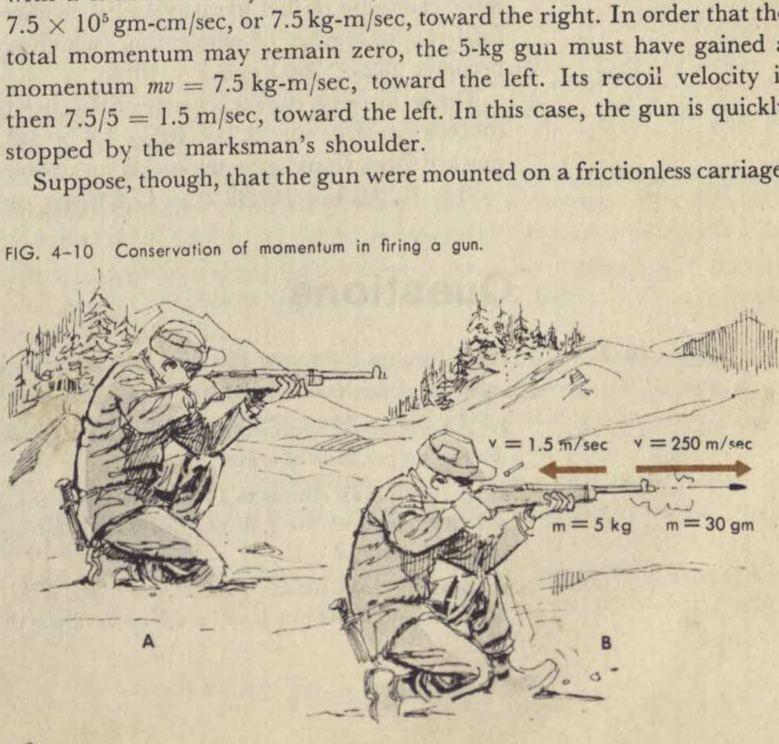
Although mechanical energy is never completely conserved in any collisions between real bodies, *momentum always is*. Two equal lumps of putty traveling at equal speeds in opposite directions will stick together and come to rest when they collide, thus using up all their KE in the work of rubbing putty particles against one another. Their total momentum, though, has not changed. Before collision, the two equal vectors representing their momentums point in opposite directions and hence add up (vectorially) to zero. After collision, the total momentum is obviously still zero.

4-8 Rocket Propulsion

As an approach to considering the great rockets that launch astronauts with tons of equipment up into their orbits, let us first consider a man firing a gun. Figure 4-10A shows the gun and bullet before firing. Both are stationary, and the total momentum is undeniably zero. A fraction of a second after firing (Fig. 4-10B), the 30-gm bullet has left the barrel with a muzzle velocity of 250 m/sec. It therefore has a momentum of 7.5×10^5 gm-cm/sec, or 7.5 kg-m/sec, toward the right. In order that the total momentum may remain zero, the 5-kg gun must have gained a momentum $mv = 7.5$ kg-m/sec, toward the left. Its recoil velocity is then $7.5/5 = 1.5$ m/sec, toward the left. In this case, the gun is quickly stopped by the marksman's shoulder.

Suppose, though, that the gun were mounted on a frictionless carriage,

FIG. 4-10 Conservation of momentum in firing a gun.



and could fire one bullet per second for a minute. At the end of the minute, after 60 shots had each given it an additional velocity of 1.5 m/sec, the gun would be moving at $1.5 \times 60 = 90$ m/sec. We would get exactly the same result if the gun (actually, a small model rocket) were able to squirt 30 gm/sec of water or hot gases or anything else out of its muzzle with a relative velocity of 250 m/sec. In the previous section, we saw that the force resulting from such a procedure is equal to the rate of change of momentum. In this example, 30 gm/sec have been given a velocity change of 250 m/sec, to give a rate of change of momentum of 250 m/sec $\times$ 0.030 kg-sec = 7.5 kg-m/sec^2. The resulting force on the rocket is thus 7.5 nt. (Note that the dimensions check: 1 nt = 1 kg-m/sec^2.) Acting on the 5-kg rocket, this force produces an acceleration $a = F/m = 7.5/5 = 1.50$ m/sec^2. After 60 sec, then, $v = at = 1.5 \times 60 = 90$ m/sec, which confirms our first result calculated from the bullet firing.

This same procedure can be used to calculate the thrust given one of the great modern satellite-launching rockets. The Titan III-C, together with its load, has a weight of 1.4×10^6 lb. Its first stage, responsible for starting the rocket on its way and for penetrating most of the atmosphere, burns solid fuel at a maximum rate of 9400 lb/sec. At sea level, the exhaust gases resulting from this burning have a velocity of 7600 ft/sec. Here, in order to calculate the thrust in *pounds*, we must have the rate of change of momentum in *slug*-feet/second^2. So we must convert the fuel-burning rate from 9400 lb/sec to $9400/g = 9400/32.2 = 292$ slugs/sec. Since each slug of fuel has its velocity change from 0 to 7600 ft/sec, the total rate of change of momentum (which equals the force, or thrust) becomes $292 \times 7600 = 2.2 \times 10^6$ lb.

The *net* force on the rocket is the difference between this thrust and the rocket's weight, or $2.2 \times 10^6 - 1.4 \times 10^6 = 8 \times 10^5$ lb. This net upward force would give the rocket an upward acceleration of $a = F/m = 8 \times 10^5/(1.4 \times 10^6/32.2) = 17.0$ ft/sec, or 0.53 g.

Questions

(4-1)

1. It requires a force of 150 lb to push a car at constant speed along a level road. How much work is done in pushing this car a quarter-mile?

2. A horse exerts a force of 250 lb in pulling a loaded wagon. How much work does he do per mile of travel?

3. A 180-lb man climbs to a roof 40 ft above the ground. (a) How much work did he do against the pull of gravity? (b) By how much did he increase his potential energy?

4. A storage tank holds 10 tons of water, and is 60 ft above the pond from which it is filled. (a) How much work is done in filling the tank? (b) How much more

potential energy does the water in the tank have than 10 tons of water in the pond?

5. In a laboratory on the ground floor of a building, a 10-lb weight is on a shelf 8 ft above the floor; a 20-lb weight is on a shelf 3 ft above the floor. (a) Which has the greater PE with reference to the floor of the laboratory? (b) Which has the greater PE with reference to the cellar floor, 10 ft below the floor of the laboratory?

6. On Planet Z there is a room whose ceiling is 10 nyorfs above the floor. A parling that weighs 7 krid is hung from the ceiling by a 2-nyorf rope; a large 40-krid parling is hung on an 8-nyorf rope. Which parling has the greater PE (a) with reference to the ceiling? (b) with reference to the floor?

(4-2)

7. How much work in joules must be done to hoist a 750-kg safe to the top of a building 35 m high?

8. A nickel coin has a mass of about 5 gm. How many ergs of work are needed to pick up a nickel from the floor and place it on a table 80 cm high?

9. An alpha particle, initially moving slowly, enters an electric field which exerts on it a force of 4×10^{-9} dyne. What is the kinetic energy of the alpha particle after having traveled 5 cm in the electric field (in ergs, and also in joules)?

10. The electric field in an X-ray tube exerts a force of 3.2×10^{-16} dyne on an electron, through a distance of 20 cm. How much kinetic energy is imparted to the electron?

11. By what factor is a car's kinetic energy increased if its speed is tripled?

12. From what height must a car be dropped in order for it to crash with as much KE as it has when traveling at 60 mi/hr?

13. How fast does a 1500-kg car have to move in order to have as much kinetic energy as a 150-gm bullet traveling 700 m/sec?

14. How fast must an alpha particle move to have as much KE as an electron traveling 8×10^6 m/sec? (An alpha particle is about 1800 times more massive than an electron.)

15. It requires a horizontal force of 2×10^4 dynes to keep a block moving at constant speed across a level table. What is the kinetic energy of the block after a horizontal force of 8×10^4 dynes has been applied to it for a distance of 1 m?

16. A push of 200 lb is applied to the car of Question 1. Through what distance must this force be exerted to give the car a speed of 8 ft/sec? (The car weighs 3200 lb.)

17. A 6080-lb truck, starting from rest, coasts down a hill whose foot is 25 ft lower than its top. When the truck reaches the bottom of the hill it has a speed of 30 ft/sec. How much energy was used up in overcoming friction on the way down?

18. A 48-lb boy slides down a slide whose bottom is 6 ft below its top. He reaches the bottom with a speed of 16 ft/sec. How much energy was used up in overcoming friction?

19. A bicyclist (total weight 160 lb) coasts 300 ft down a hill sloping at 10°; the road then climbs up the opposite side of the valley at a 10° slope, and he

coasts 100 ft up this road before he comes to a halt. What was the average force of friction?

20. A 4800-lb truck coasts 200 ft down a 12° hill; the road then changes to an uphill 8°, and the truck coasts 150 ft along this uphill slope before stopping. What was the average force of friction?

(4-3)

21. A pendulum is pulled to one side until its center of gravity is 3 cm higher than when hanging vertical, and is then released. What is its speed as it passes through the midpoint of its swing?

22. A pendulum is pulled to one side until its center of gravity is 3 inches higher than when hanging vertical, and then released. What is its speed as it passes through the midpoint of its swing?

(4-4)

23. A horizontal pipe of 10 cm^2 cross-sectional area is smoothly reduced to a cross-sectional area of 5 cm^2. The pipe carries water, which flows through the larger section with a speed of 30 cm/sec. What is the difference in pressure between the large and small sections?

24. Oil of density 0.86 gm/cm^3 flows through a horizontal pipe at a speed of 50 cm/sec. The pipe (cross-sectional area 30 cm^2) is smoothly narrowed to a cross-sectional area of 10 cm^2, and pressure gauges are placed in the large and the small sections of pipe. What is the difference in their pressure readings?

25. Oil of density 0.90 gm/cm^3 flows through a pipe whose cross-sectional area is 30 cm^2. This pipe is smoothly reduced to a cross-sectional area of 10 cm^2. Gauges in the large and small pipe sections show a pressure difference of 3000 dynes/cm^2. How many cm^3 of oil flow through the pipe per second?

26. A horizontal pipe of 10 cm^2 cross-sectional area is smoothly reduced in size to a cross-sectional area of 5 cm^2. A gauge shows that the difference in pressure between the large and small sections of pipe is 1200 dynes/cm^2. How many cm^3/sec of water are flowing through the pipe?

(4-5)

27. The man in Question 3 makes the climb in 50 sec. What power is he exerting? (In ft-lb/sec, and in horsepower.)

28. The tank in Question 4 is to be filled in 40 min. What is the power that the pump must deliver? (In ft-lb/sec, and in horsepower.)

29. A motor and construction hoist mechanism are 60 percent efficient. What is the minimum horsepower of the motor needed to raise a load of $2\frac{1}{2}$ tons to a height of 120 feet in 2 minutes?

30. A ski lift must raise a total load of 2500 lb each minute up a slope whose difference in elevation is 1500 ft. If the mechanism is 70 percent efficient, what is the horsepower of the engine?

31. A tank on the roof of a building 20 m high holds 10 m^3 of water, and it must be filled from a pond on the ground in 20 min. What is the power in watts of the pump that will be required, assuming it to be 100 percent efficient? 60 percent efficient?

32. A load of 800 kg is to be hoisted to an elevation of 35 m in 3 minutes. What is the power in watts of the driving motor, assuming it to be 100 percent efficient? 70 percent efficient?

33. A 200-horsepower engine can give an airplane a speed of 120 mi/hr in level flight. What is the force of friction at this speed?

34. What is the retarding force of friction on a car which must exert 150 horsepower to drive it at 80 mi/hr on a level road?

35. How many joules of electrical energy are converted into heat and light when a 50-watt lamp burns for 3 hr?

36. A 250-watt heater burns for 6 hr. How many joules of energy are converted from electrical energy into heat?

(4-7)

37. How fast does a 1500-kg car have to move in order to have as much momentum as a 150-gm bullet traveling 700 m/sec? (Refer to Question 13, and compare answers.)

38. How fast must an alpha particle move in order to have the same momentum as an electron traveling 8×10^6 m/sec? (Refer to Question 14, and compare answers.)

39. For how long must a push of 10,000 newtons be applied to a 50,000-kg rocket ship in order to increase its speed from 25,000 m/sec to 30,000 m/sec?

40. An 8-lb block at rest on a level, frictionless table is given a constant 8-oz push until it is moving 2 ft/sec. For how long a time was the push applied?

41. An alpha particle (mass $= 6.6 \times 10^{-24}$ gm) with a velocity of 3×10^7 cm/sec is subjected to a force of 10^{-9} dyne in the direction of its motion for 10^{-6} sec. What is the resulting velocity of the particle?

42. A 16-ton rocket ship traveling freely in space at 6000 mi/hr is given a thrust of 5000 lb in the direction of its motion for 5 minutes. What is its velocity at the end of this thrust?

43. A 2000-lb car traveling 60 mi/hr collides with a 6000-lb truck going in the same direction at 30 mi/hr. The bumpers lock together. What is the speed of the locked-together cars immediately after the collision?

44. A 400-gm glider on an air track is moving 10 cm/sec when it is overtaken by a 600-gm glider moving at 15 cm/sec. The two gliders lock together. What is their speed after this collision?

45. A proton moving east with a velocity of 10^5 cm/sec scores a direct hit on a stationary alpha particle. (The mass of an alpha particle is 4 times the mass of a proton.) The collision is perfectly elastic. What are the velocities (both speed and direction) of the two particles after collision?

46. An alpha particle moving east with a velocity of 10^5 cm/sec scores a direct hit on a stationary proton. (The mass of an alpha particle is 4 times the mass of a proton.) The collision is perfectly elastic. What are the velocities (speed and direction) of the two particles after collision?

47. A bullet weighs 10 gm and has a speed of 300 m/sec. It is fired into a stationary 5-kg block of wood on level ice, and remains imbedded in the wood. (a) What is the velocity of the block after the impact? (b) What fraction of the KE of the bullet has been converted into frictional heat?

48. A 160-lb man sits on a level frictionless surface. A friend throws him a 1-lb ball at a speed of 90 ft/sec, which he catches and holds. (a) What is the speed of the seated man after catching the ball? (b) What fraction of the KE of the ball has been converted into frictional heat?

(4-8) **49.** Suppose the rocket in Question 39 has an exhaust velocity of 2500 m/sec. At what rate must it burn fuel to develop its thrust of 10^4 newtons?

50. A 20-kg model rocket has an exhaust velocity of 2000 m/sec. At what rate must it burn fuel to have an acceleration of 1 g?

51. A fire hose is 3 inches in diameter, and its nozzle is 1 inch in diameter. If the speed of the water through the hose is 10 ft/sec, with what force must the nozzle be held to prevent its flying backward?

52. A hose has a diameter of 8 cm, and delivers a jet of water through a nozzle 2 cm in diameter. Water flows through the hose at a speed of 1 m/sec. What is the backward force on the nozzle?

53. With what force would a 200-lb astronaut press against his couch in taking off in a Titan III-C rocket launch?

54. If the rocket ship of Question 39 were in open space, away from perceptible gravitational forces, what would be the apparent "weight" of an astronaut aboard it during the described acceleration? The astronaut weighs 175 lb on the earth.

chapter / five

Rotational Mechanics

5-1 Equations of Rotational Motion

There are two ways in which a body can move. One type of movement, which we have already investigated, is called *linear*, or *translational*, and is a mere change in position. The other type is called *angular*, or *rotational*, movement, in which the body rotates about a line called an *axis*. Many motions, of course, are a combination of the two; a spiral punt spins as it soars through the air, and the wheels of a car rotate as they travel along the road.

In translational motion, our basic measure was the distance traveled; in rotation, we must start with the angle through which the body turns. It might seem good enough to measure this angle in degrees, or in terms of complete revolutions, but in practice the relationships and equations are much simpler and easier to comprehend if we use another unit called the *radian* (rad).

Figure 5-1A shows what a radian is. Take a circle of any size and lay off along its circumference an arc equal in length to the radius of the circle. This arc measures an angle of 1 radian. Since the circumference

is $2\pi r$, it is apparent that there are 2π radians in a full circle, or revolution.

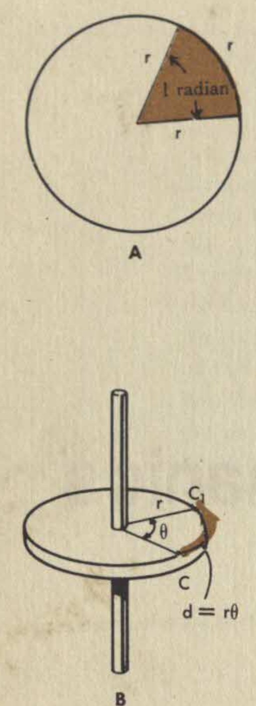


FIG. 5-1 The radian (A), and a rotating disk (B).

Imagine a wheel of radius r, with a point C marked on its rim (Fig. 5-1B). If we turn the wheel on its axis through an angle of θ radians, point C moves in an arc through a distance d to a new position, C_1. If θ is 1 radian, the distance C–C_1 will equal r; if θ is 2 radians, C–C_1 will be $2r$. In general, $d = \theta r$; that is, the distance moved by any point on a rotating body is the angle (in *radians*!) through which the body turns, multiplied by the distance from the point to the axis of rotation.

Linear velocity and angular velocity are similarly related. If our wheel rotates at a uniform rate through θ radians in t seconds, a point at distance r from the axis will have moved through a distance θr in the same time t. The angular speed ω is θ/t rad/sec, and the linear velocity of point C is $v = \theta r/t$. From this we see that $v = \omega r$. Further analysis will show that for a point at a distance r from the axis of a wheel whose angular speed is changing, the linear acceleration of the point along its curved path is r times the angular acceleration of the wheel, measured in rad/sec^2. Angular acceleration α is of course defined as the rate at which ω changes, or $\Delta\omega/t$. Stated as an equation, $a = \alpha r$. Here are these three simple relationships tabulated:

$$d = r\theta \qquad v = r\omega \qquad a = r\alpha.$$

By starting from basic definitions of angular velocity and angular acceleration, we can easily derive equations of rotational motion exactly as they were derived for translational motion:

$$\omega_t = \omega_o + \alpha t$$
$$\theta = \omega_o t + \tfrac{1}{2}\alpha t^2$$
$$\omega_t^2 = \omega_o^2 + 2\alpha\theta$$

5-2 Moment of Inertia

An applied force causes a mass to accelerate and move from one location to another. In the rotational analogy, an applied torque causes a body to rotate, and it seems apparent that there must be some properties of a rotating body we can use to get a relationship corresponding to $F = ma$. Let us investigate this by considering the simplest possible rotating body, as shown in Fig. 5-2. It consists of a rod (imagined to be massless) extending from an axis of rotation O to a small mass m, at a distance r from O. Now, to the mass m apply a force F, as shown. The torque tending to rotate the body around O is Fr. From Newton's second law, a force F acting on a mass m will give it a linear acceleration of $a = F/m$. We know also that linear acceleration is related to angular acceleration by the expression $a = r\alpha$. This gives two expressions for a, which we can set equal to one another:

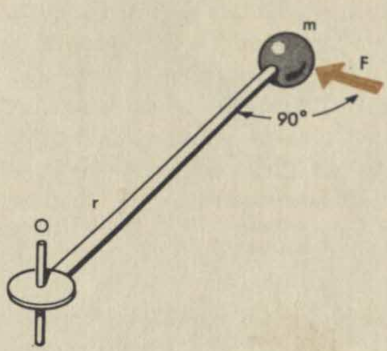


FIG. 5-2 A force applied to a small mass at the end of a massless rod causes it to rotate about the axis at O.

$$\frac{F}{m} = r\alpha$$

from which

$$F = mr\alpha.$$

For rotational motion we need something including the torque τ instead of force. We can get this simply by multiplying the equation by r:

$$Fr = mr^2\alpha$$

or

$$\tau = mr^2\alpha.$$

The result is an equation exactly analogous to $F = ma$, except that instead of m we have mr^2. This expression, the mass of a particle times the square of its distance from the axis of rotation, is called the particle's *moment of inertia* and is usually designated by the letter I. Our rotational form of Newton's second law thus becomes

$$\tau = I\alpha.$$

We have worked this expression out for a single small particle, but what of the rods and wheels and cylinders we are most often concerned with when we need to consider rotation? Any large body can be thought of as being made up of a large number of small particles. Each particle contributes its own small mr^2, so the moment of inertia of the large body is the sum of the moments of inertia of the particles that make it up. For regularly shaped bodies of many sorts, the branch of mathematics known as *calculus* gives these sums, some of which are shown in Fig. 5-3.

Without calculus, the only one of these we can satisfactorily confirm is the thin-walled hollow cylinder. If we consider the cylinder to be made of a multitude of small pieces of mass m, each small piece is the same distance R from the axis of rotation, and each has the same moment

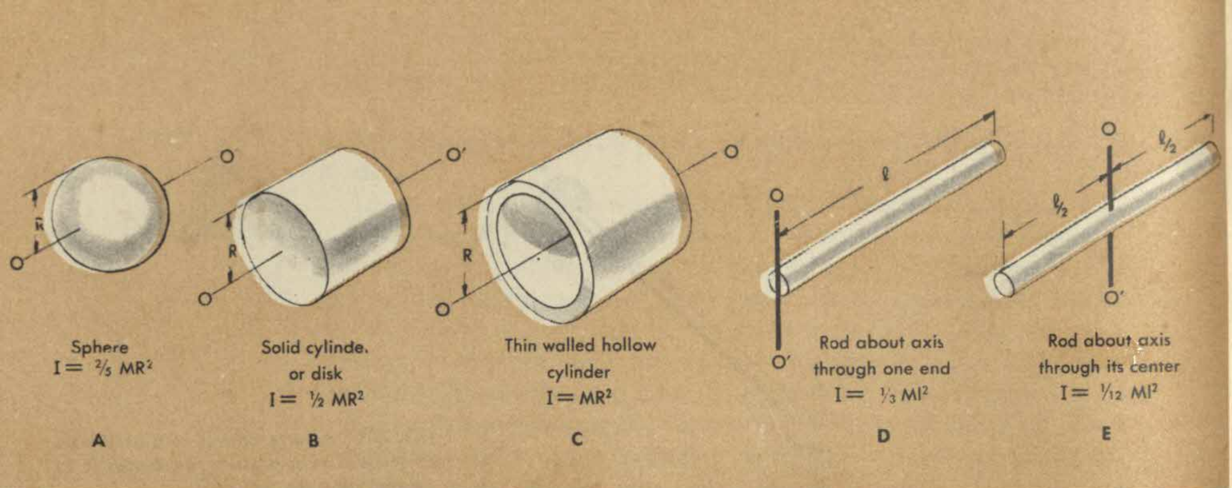


FIG. 5-3 Moments inertia of bodies of various shapes.

of inertia mR^2. If all these small moments of inertia are added together, the sum is clearly R^2 times the sum of the small masses, or MR^2, where M is the total mass of the cylinder.

5-3 Torque and Rotation

As an example of how to use moment of inertia in the rotational form of Newton's second law, consider a solid cylinder on a frictionless axle, similar to that shown in Fig. 5-3B. The weight of the cylinder is 128 lb, and it is 6 inches in *diameter*. A steady pull of 24 lb is exerted on a cord wrapped around the cylinder (Fig. 5-4). Starting from rest, how long a time will it take the cylinder to make 12 revolutions? We can first determine its angular acceleration from $\tau = I\alpha$, keeping a careful eye on units. If we use torque in lb-ft (*not* lb-in., because the distance unit must be the same as the distance unit in g, which we will take as 32 ft/sec²), then the mass must be taken in slugs. This gives us

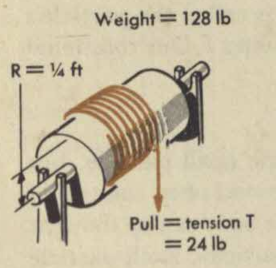


FIG. 5-4 Rotating cylinder given an angular acceleration by a 24-pound pull on a cord wrapped around it.

$$\tau = I\alpha = \tfrac{1}{2}MR^2\alpha$$

$$24 \times \tfrac{1}{4} = \tfrac{1}{2} \times \tfrac{128}{32} \times \tfrac{1}{16} \times \alpha$$

$$\alpha = 48 \text{ rad/sec}^2.$$

Once we have α, the rest is easy, for 12 revolutions is $12 \times 2\pi = 24$ radians, and from $\theta = \frac{1}{2}\alpha t^2$,

$$24\pi = \tfrac{1}{2} \times 48t^2$$

$$t^2 = \pi$$

$$t = 1.77 \text{ sec.}$$

At first glance, it might appear that in the last example we could get the steady 24-lb pull we want by merely hanging a 24-lb weight on the end of the cord. A closer look, however, will convince us that this is not true. Figure 5-5 sketches the situation. If the axle of the cylinder is

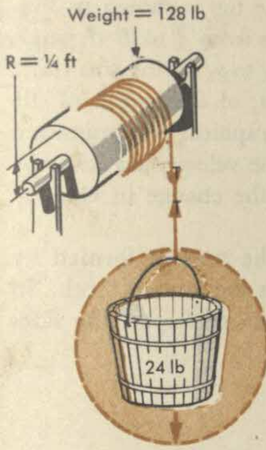


FIG. 5-5 Rotating cylinder given an angular acceleration by a 24-pound weight hung from a cord wrapped around it.

clamped so that it cannot turn, the weight will be motionless and the cord tension T will be 24 lb. But if the cylinder is free to turn on frictionless bearings, it will begin an angular acceleration, and the weight will have a linear acceleration downward. Let us look at just the descending weight, which is isolated in the dashed circle. It has a downward acceleration a, caused by the net force $24 - T$, so the cord tension must be less than 24 lb. $F = ma$ gives us

$$24 - T = \frac{24}{32}a = \frac{3a}{4}.$$

The torque of force T is what causes the cylinder to have an angular acceleration, and we can write another equation:

$$\tau = I\alpha = \tfrac{1}{2}MR^2\alpha$$

$$\tfrac{1}{4}T = \tfrac{1}{2} \times \tfrac{128}{32} \times (\tfrac{1}{4})^2\alpha = \tfrac{1}{8}\alpha$$

from which

$$\alpha = 2T.$$

We cannot solve two equations for three unknown quantities, but fortunately we know that $a = r\alpha$, which gives another relation between the unknowns:

$$a = \tfrac{1}{4}\alpha$$

or

$$\alpha = 4a.$$

Substitution into the equation $\alpha = 2T$ gives us

$$4a = 2T$$

$$T = 2a.$$

Another substitution into the first equation gives

$$24 - 2a = \tfrac{3}{4}a$$

from which

$$a = 8.73 \text{ ft/sec}^2$$

and

$$T = 17.5 \text{ lb.}$$

5-4 Centripetal and Centrifugal Force

Velocity, we have noted before, is a vector quantity. So far, we have considered only accelerations caused by a change in speed, but a vector can also be changed by changing its direction.

Figure 5-6 shows such a case. It represents a mass m at the end of a massless rod of length r, which rotates at constant speed about the axis

OO' (Fig. 5-6A). A view looking down from the top is shown in Fig. 5-6B. Let us say that in a short time t, m moves from A to B through an angle θ. Its velocity will have changed from v_1 to v_2 during this time, a change in *direction* only; the speed is constant, so the length of the vector remains the same. Figure 5-6C shows a graphical determination of the change in velocity that has occurred. The velocity Δv must be added to v_1 to produce v_2; thus Δv represents the change in velocity during time t.

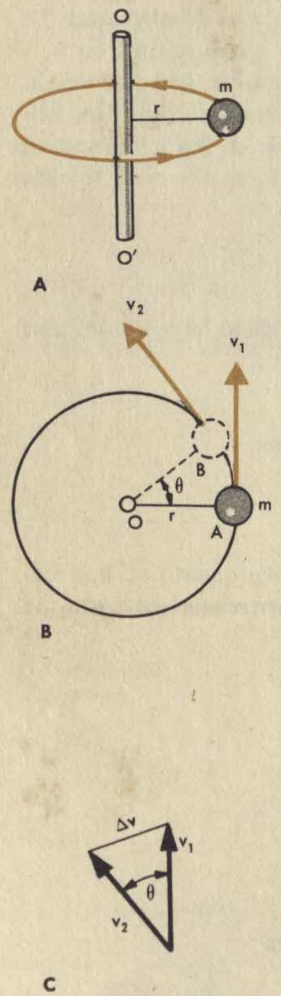


FIG. 5-6 Centripetal acceleration of a body moving in a circle at constant speed.

Notice that the triangle OAB is similar to the triangle formed by the vectors v_1, v_2, and Δv, because the angle θ is the same in both. In similar triangles (i.e., triangles of the same shape), corresponding sides are in the same proportions, so we can write

$$\frac{\Delta v}{v} = \frac{AB}{r} = \frac{vt}{r}$$

$$\Delta v = \frac{v^2 t}{r}.$$

(The subscript has been dropped from the v because we are now dealing with magnitudes only; since v_1 and v_2 are the same length, we do not need the subscripts to distinguish them.)

Acceleration a we have defined as the rate of change of velocity. The change Δv has occurred in the time t, so its *rate* of change is

$$a = \frac{\Delta v}{t} = \frac{v^2}{r}.$$

In some problems, it may be more convenient to use the angular speed ω rather than the linear speed v. We can make this change easily by substituting ωr for v, giving us

$$a = \omega^2 r.$$

Since a is a vector, we must also determine its direction. If, in Fig. 5-6C, we make the time interval, and hence also the angle θ, smaller and smaller, v_1 and v_2 become practically parallel, and Δv becomes perpendicular to either velocity (or to both). Thus Δv, and accordingly a also, point directly inward toward the center of the circle.

There can be no acceleration without a force to cause it, and we see at once that the center-directed acceleration of our mass is caused by the constant inward pull of the rod on which it is mounted. The force that the rod exerts on the mass is called the *centripetal* force; there is, of course, an equal and opposite *centrifugal* force that the mass exerts on the rod. The magnitude of these forces comes from our old friend:

$$F = ma \qquad F = \frac{mv^2}{r} \quad \text{or} \quad F = mr\omega^2.$$

As an example, let us consider a small boy who finds in his pocket

a doorknob tied to a 60-cm length of string. Naturally, he starts whirling the doorknob in a circle in a vertical plane. How many revolutions/minute must he whirl it to keep the string from becoming slack at the top of the circle? We do not know the mass of the knob, but to put something in our equations we can call the mass m gm. At any point on the circle, the knob will need an inward pull of $mr\omega^2 = 60m\omega^2$ to keep it in its circular path. At the top of the circle, the pull of gravity mg is acting directly toward the center; so if mg is greater than the required centripetal force $mr\omega^2$, the knob will be pulled inward from its circular path and the string will become slack. The lowest possible speed that will prevent this slackening of the string is when

$$mg = mr\omega^2$$

$$g = r\omega^2$$

$$\omega^2 = \frac{g}{r} = \frac{980}{60}$$

$$\omega = 4.04 \text{ rad/sec.}$$

We know that 2π rad = 1 rev and that 60 sec = 1 min, so to convert our figure to rev/min, we write

$$\frac{4.04 \text{ rad}}{\text{sec}} \times \frac{1 \text{ rev}}{2\pi \text{ rad}} \times \frac{60 \text{ sec}}{1 \text{ min}} = 38.6 \frac{\text{rev}}{\text{min}}.$$

It is worth noticing that m, the mass of the doorknob, cancels out of the equations, so that our answer is the same for an object of any mass.

5-5 Rotational Work and Energy

Because work is required to make a heavy wheel spin and because the spinning wheel can be made to do work, it is apparent that kinetic energy is stored in a rotating body in much the same way that it is stored in a body having translational motion. To calculate the KE of a rotating wheel, we need only add the KE's of all the small particles that we consider the wheel to be made up of. Look back to Fig. 5-2 and imagine the mass m to be one small particle of a wheel revolving about axis OO'. If this small particle is moving at a linear speed v, its KE is $\frac{1}{2}mv^2$.

It would be very difficult, however, to add up the KE's of all the particles of the wheel if they were expressed in terms of v, because the speed of each particle depends on its distance from the wheel's axis of rotation.

We can get around this difficulty by converting our expression for KE into one in terms of ω, because the *angular* speed is the same for all the particles of a rotating body:

$$v = r\omega, \quad \text{so} \quad v^2 = r^2\omega^2.$$

Then

$$KE = \tfrac{1}{2}mv^2$$

becomes

$$KE = \tfrac{1}{2}mr^2\omega^2 = \tfrac{1}{2}I\omega^2.$$

Here the I is the moment of inertia of the whole rotating body, as given in Fig. 5-3, and we have an exact analog of the KE $= \frac{1}{2}mv^2$ used in translational motion.

In investigating the mechanics of translation, we saw that a body could be given energy—either potential or kinetic—by doing work on it. The same ideas apply to rotation. Of course, a body cannot be given gravitational potential energy merely by rotating it, but there are other forms of potential energy. You do work when you wind your watch, and this is stored as PE in the wound-up spring. Work done in rotation is measured by the analog of force $\times$ distance—that is, torque $\times$ angle:

$$W = \tau\theta \quad (\theta \text{ in radians, naturally!}).$$

To give a moving body a certain amount of kinetic energy, we must do an equal amount of work on the body: $Fd = \frac{1}{2}mv^2$. Similarly, to give a rotating wheel, for example, a certain amount of rotational kinetic energy, we must also do an equal amount of work: $\tau\theta = \frac{1}{2}I\omega^2$.

Let us use the idea of rotational kinetic energy to compare the speeds with which a hoop and a solid disk reach the bottom of a slope. Figure 5-7 shows the hoop and disk ready to start their trip down, a journey which will consist of both rotation and translation. If no energy is lost in friction, the KE at the bottom of the slope must equal the PE at the top, or

$$\tfrac{1}{2}Mv^2 + \tfrac{1}{2}I\omega^2 = Mgh.$$

For the hoop,

$$I = MR^2$$

so

$$\tfrac{1}{2}Mv^2 + \tfrac{1}{2}MR^2\omega^2 = Mgh.$$

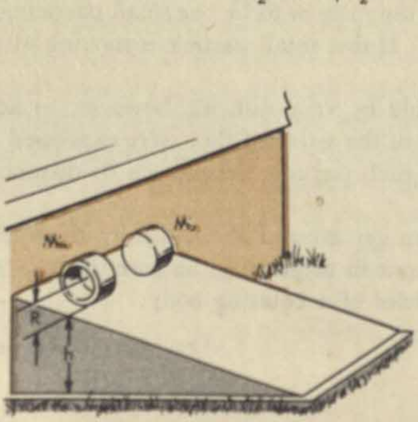


FIG. 5-7 Hoop and solid cylinder ready to race down an inclined plane.

The mass of the hoop cancels out, which tells us that whether it is a wooden hoop or a lead hoop of the same size makes no difference, so that

$$\tfrac{1}{2}v^2 + \tfrac{1}{2}R^2\omega^2 = gh.$$

We are not through yet, since we know that $v = \omega R$. Thus $R^2\omega^2 = v^2$, and the equation becomes

$$v_H^2 = gh \quad \text{or} \quad v_H = \sqrt{gh}.$$

We see now that R has vanished, as the mass did, so the size of the hoop has no effect on the speed. A hoop made of paper, a hoop made of iron, a small hoop, and a large hoop will all roll down the slope side by side.

If you do the same computation for the solid disk, whose moment of inertia is $\tfrac{1}{2}M_DR^2$, you will find that

$$v_D = \sqrt{\tfrac{4}{3}gh}.$$

Therefore, the solid disk, which travels faster than the hoop, will reach the bottom first. How will a rectangular block do in this race, if it slides down the slope without friction?

5-6 Angular Momentum

Just as linear momentum was defined as mv, so *angular momentum*, which is the momentum a body has because of its rotation, is defined as $I\omega$. Newton's third law of action and reaction applies to torques as well as to forces, and the law of conservation of angular momentum can be proved in the same way as conservation of linear momentum.

The conservation of angular momentum can be very easily demonstrated by a student standing (or sitting) on a nearly frictionless rotating platform (Fig. 5-8). Someone sets her spinning at a moderate speed which will presumably stay constant until it is gradually diminished by the unavoidable torque of friction. But if the passenger extends her arms (Fig. 5-8B), she at once slows down; when she returns her arms to their original position, she regains her former speed (Fig. 5-8C). The principle of conservation of angular momentum tells us that (neglecting the effect of the torque of bearing friction) the product $I\omega$ remains unchanged. So, when the rotating student increases her moment of inertia by extending her arms, her angular speed diminishes in the same proportion; when she reduces I, ω must immediately increase to keep the product $I\omega$ constant.

Spaceships will occasionally need to reverse the direction in which they point in flight—when, for instance, they want to come in for a landing on Mars tail-first with their rockets blasting in order to brake the ship to a gentle stop. It has been suggested that this could be done by mounting a heavy wheel on an axis perpendicular to the axis of the ship. If the wheel is set rotating, the ship must rotate in the opposite direction so that the total angular momentum will continue to add up to

zero. (The ship will turn much more slowly than the wheel, of course; since its I is enormously larger than that of the wheel, its ω must be correspondingly smaller.) When the ship has rotated far enough, one need only stop the wheel and the ship must also stop rotating.

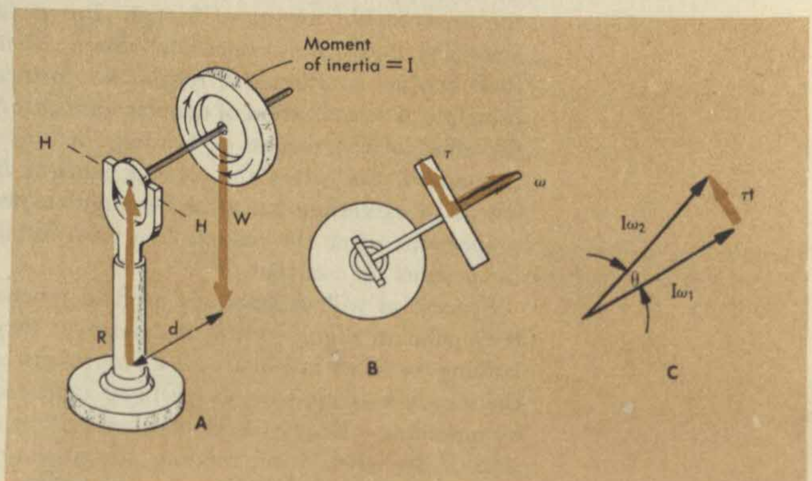
The behavior of a gyroscope is quite puzzling if we try to apply to it the rules of translational forces and motions we are most familiar with. A gyroscope is merely a heavy wheel that can be set spinning at high speed. It is generally mounted in low-friction bearings on a frame so that it can be moved and handled while spinning. Figure 5-9A shows a rapidly spinning wheel with its axle supported at one end in a bearing that is free to rotate about the stand and can also pivot freely about the horizontal axis *HH*. The curved arrows show the direction of rotation of the wheel; *W* and *R* represent the weight of the gyroscope and the equal upward reaction of the stand that supports it.

Torque, angular velocity, and angular momentum can all be represented by vectors, according to what is called the "right-hand screw rule." Imagine turning an ordinary screw in the direction the wheel is rotating. The threads of the screw would cause it to advance along the axle toward the unsupported end. The angular velocity of the wheel is represented by a vector ω in this direction, as is indicated in the top view shown in Fig. 5-9B. The torque Wd is similarly shown by the vector τ.

At first glance, we would expect the torque τ to cause the unsupported end of the axle to fall, as would certainly be the case if it were acting on a brick or a ruler or some other nonrotating body. A closer look, however, will reveal that the rapid spinning of the wheel makes it behave

FIG. 5–8 Demonstration that the angular momentum of a rotating body remains constant when no external torque is applied.

FIG. 5–9 Gyroscopic precession.



quite differently. Since I is a scalar, the vector representing the angular momentum $I\omega$ will be in the same direction as ω. Figure 5-9C shows $I\omega_1$ pointing along the axle at some particular instant. Let us find what effect the torque will have by applying τ for a short period of time t. This will give the wheel an impulse τt (analogous to a linear impulse Ft) in the direction of τ. We now recall that impulse equals change in momentum, and we add the change τt vectorially to the original momentum $I\omega_1$ to get the new momentum vector $I\omega_2$, which represents the wheel's momentum, and consequently the direction of the axle, at the end of time t. The angle θ through which the axle has turned is $\tau t/I\omega$ radians, and it has turned through this angle in t seconds. Thus it has swung around at an angular speed Ω, called *precession*, equal to θ/t, or $\tau/I\omega$. This gives us the following fundamental gyroscopic equation:

$$\Omega = \frac{\tau}{I\omega}.$$

The torque does *not* cause the spinning wheel to drop, but to swing around (in our example) in a counterclockwise direction, viewed from the top.

If we assume the wheel, which has a heavy rim, to be almost a hoop, we can figure its moment of inertia as MR^2. For a 500-gm wheel 5 cm in radius, I would be 12,500 gm-cm^2. If the distance d is 4 cm, τ is $4 \times 500 \times 980 = 1{,}960{,}000$ dyne-cm. A reasonable rotational speed would be 1800 rev/min, or $1800 \times 2\pi/60 = 188$ rad/sec. Under these conditions, we could confidently predict that the wheel would precess at the rate of $1.96 \times 10^6/(1.25 \times 10^4 \times 1.88 \times 10^2) = 0.835$ rad/sec.

Questions

(5-1)

1. A wheel has 12 equally spaced spokes. What is the angle between adjacent spokes, in radians?

2. Through how many radians does the earth rotate between noon at Greenwich (0° latitude), and noon at Denver (105° W. lat.)?

3. (a) What is the angular speed of the earth's rotation about its axis, in rad/hr? (b) With this figure determine the linear speed of a point on the earth's equator, using 8000 miles as the earth's *diameter*.

4. (a) What is the angular speed of the earth's revolution around the sun, in rad/hr? (b) With this figure, determine the earth's orbital speed, taking the radius of its orbit to be 93×10^6 mi.

5. A grinding wheel 8 inches in diameter rotates with an angular speed of 2400 rev/min. (a) Convert this angular speed to rad/sec. (b) What is the linear speed of the rim of the wheel, in ft/sec?

6. A centrifuge 10 cm in diameter rotates at 30,000 rev/min. (a) Convert this angular speed to rad/sec. (b) What is the linear speed of the rim of the centrifuge?

7. It requires the wheel of Question 5 four sec to obtain its final speed. What is its average angular acceleration, in rad/sec^2?

8. The centrifuge of Question 6 takes 4 min to attain its final speed. What is its average angular acceleration, in rad/sec^2?

9. A wheel accelerates uniformly at 60 rad/sec^2. Through how many revolutions does it turn to attain a speed of 3000 rev/min, starting from rest?

10. A dental drill accelerates from rest to 9000 rev/min in 2 sec. (a) What is its average angular acceleration? (b) How many revolutions does it make in coming up to full speed?

(5-2) **11.** What is the moment of inertia of a 10-kg sphere 30 cm in diameter?

12. What is the moment of inertia of a 160-lb cylinder 2 ft in diameter?

13. A slender rod 2 m long has a mass of 8 kg. What is its moment of inertia about its center?

14. What is the moment of inertia of the rod of Question 13 about an axis at one end?

(5-3) **15.** A cylindrical grinding wheel has a mass of 3 kg and is 20 cm in diameter. (a) What is its moment of inertia? (b) What torque will be required to give it an angular acceleration of 120 rad/sec^2?

16. A circular saw (which can be considered to be a very short solid cylinder) is 8 inches in diameter, and weighs 4 lb. What torque is required to give it an angular acceleration of 80 rad/sec^2?

17. It is found that a torque of 8 oz-in. gives an irregularly shaped wheel an acceleration of 2 rad/sec^2. What is the moment of inertia of the wheel?

18. An irregular 5-kg wheel is given an acceleration of 10 rad/sec^2 by an applied torque of 2×10^7 dyne-cm. What is the moment of inertia of the wheel?

19. Refer to Fig. 5-4. What weight would have to be suspended from the cord in order to produce the 24-lb tension shown?

20. What weight would have to be suspended from the cord of Fig. 5-4 in order to produce a 40-lb tension in it?

(5-4) **21.** A 50-ton locomotive goes at 45 mi/hr around a curve whose radius is 800 ft. With what force do the rails press against the flanges of the wheels?

22. A 2560-lb car rounds a level 300-ft radius curve at 40 mi/hr. With what frictional force must the paving press the tires toward the inside of the curve?

23. The mass of the earth is very nearly 6×10^{27} gm, and the radius of its orbit around the sun is about 1.50×10^{13} cm. (a) What is the force of gravitational attraction between earth and sun? (b) How much is this, in metric tons?

24. The mass of the moon is about 7.4×10^{25} gm, and the radius of its orbit around the earth is about 3.8×10^5 km. It makes one revolution around the earth in 27.3 days. (a) What is the force of gravitational attraction between earth and moon? (b) How much is this, in metric tons?

25. A small mass hangs from a thread 20 cm long. The mass is struck and given an initial velocity v which will carry the mass around in a vertical circle just fast enough so that the string will not become limp at the top of the circle. What is the required value of v? Let us lead up to this with a few intermediate questions: (a) What will be the necessary speed v_T at the top of the circle?

(b) What is the KE of the mass at the top? (c) its PE? (d) What, then, must be the KE + PE at the bottom of the circle? (e) What is the PE at the bottom? (f) What KE must it have at the bottom? (g) What speed must it have at the bottom?

26. On Planet Z, where $g = 6$ m/sec^2, a small mass hangs from a thread 30 cm long. The mass is struck and given an initial velocity v which will carry the mass around in a vertical circle just fast enough so that the string will not become limp at the top of the circle. What is the required value of v? (See Question 25, if necessary.)

27. A small carousel is 4 m in diameter and rotates once in 10 sec. If we place a wood block on the floor at the outer edge of the carousel, will it stay in place or be flung off if its coefficient of friction is 0.10?

28. Patrons of a certain amusement park concession stand with their backs against the wall of a cylindrical room 12 ft in diameter. The room rotates about its central axis with increasing speed until a certain angular velocity is reached. The operator then lowers the floor, and the patrons remain in position against the wall, held there by friction. What angular velocity must the room have, if $\mu = 0.30$?

(5-5)

29. A 10-kg cylinder 12 cm in diameter rotates about its axis at an angular speed of 2 rev/sec. What is its rotational KE?

30. A sphere 10 cm in diameter has a mass of 3 kg, and rotates at 5 rev/sec. What is its rotational KE?

31. What will be the velocity of a sliding frictionless block, in terms of the height of a slope down which it slides?

32. What will be the velocity of a rolling sphere, in terms of the height of a slope down which it rolls?

33. A hoop rolls down an inclined plane without slipping. What fraction of its KE is translational? rotational?

34. A solid cylinder rolls down an inclined plane without slipping. What fraction of its KE is translational? rotational?

(5-6)

35. A student standing on a ball-bearing platform with his arms at his sides has an $I = 2$ slug-ft^2. He is given an angular velocity of 1 rev/sec; he then extends his arms out horizontally, increasing his I to 4 slug-ft^2. What is his angular velocity with arms extended?

36. A student with arms extended horizontally stands on a rotating platform, and is given an angular speed of 1 rev in 1.5 sec. In this position, his I is 3 slug-ft^2. He brings his arms straight down at his sides, and speeds up to 1.5 rev/sec. What is his I in his latter position?

37. Compare the student's rotational KE in his two positions of Question 35. If there is a change in KE, how do you account for it?

38. In Question 36, does the student have the same KE in both positions? If not, explain.

39. Two wheels are mounted side by side on the same axle. Wheel A, whose moment of inertia is 5×10^5 gm-cm^2, is set spinning at 600 rev/min. Wheel B, with a moment of inertia of 2×10^6 gm-cm^2, is stationary. A clutch now acts

to join A and B so that they must spin together. (a) At what speed will they rotate? (b) Suppose the clutch acts gradually. Will the end result be the same as though they were joined suddenly? (Assume the bearings are frictionless.) (c) How does the rotational KE before joining compare with the KE afterward? (d) What torque will the clutch have had to transmit if A makes 10 revolutions relative to B during the operation of the clutch?

40. A spaceship has a moment of inertia about its center of 8×10^8 kg-m^2. The direction in which it points is controlled by a wheel with a heavy rim having a mass of 100 kg and a mean radius of 50 cm. How long will the wheel have to be rotated at 600 rev/min in order to rotate the ship by 90°? (Note that this can have no effect on the direction in which the ship travels.)

41. If the crankshaft and flywheel of a car rotate clockwise when viewed from the driver's seat, will the car tend to nose down or nose up when the car makes a right turn?

42. Some planes (now obsolete) had massive engines which rotated counterclockwise as viewed by the pilot sitting behind them. When the plane nosed down to go into a dive, would it swing to the left or to the right as a result of gyroscopic action?

chapter / six

Orbits and Satellites

6-1 Newton's Law of Gravitation

It is interesting to apply the laws of centripetal force to the moon as it swings around the earth in a nearly circular orbit with a radius averaging 384,000 km. The moon takes 27.3 days to make one revolution about the earth, which in radians/second is

$$\omega = \frac{1\ \text{rev}}{27.3\ \text{day}} \times \frac{2\pi\ \text{rad}}{1\ \text{rev}} \times \frac{1\ \text{day}}{24\ \text{hr}} \times \frac{1\ \text{hr}}{3600\ \text{sec}}$$

$$= 2.67 \times 10^{-6}\ \text{rad/sec.}$$

We also know that

$$r = 384{,}000\ \text{km}$$
$$= 3.84 \times 10^{10}\ \text{cm.}$$

This gives us for the moon's centripetal acceleration toward the earth

$$a = r\omega^2$$
$$= 3.84 \times 10^{10} \times (2.67 \times 10^{-6})^2$$
$$= 0.273\ \text{cm/sec}^2.$$

In the latter part of the seventeenth century, Newton did calculations similar to the ones above and made the proposal (very radical and controversial at the time) that the moon's earthward centripetal acceleration is caused by the same force that causes an apple to accelerate earthward when falling from a tree. The acceleration of the falling apple—g, or about 980 cm/sec^2—is much greater than the moon's earthward acceleration that we have just figured, but it seemed reasonable to Newton that the earth's force of gravitational attraction should become weaker with increasing distance. In fact, the figures can be quite satisfactorily reconciled by assuming that the gravitational pull decreases in proportion to the square of the distance from the earth's center; i.e., at twice the distance, the force is $\frac{1}{4}$ as great; at 3 times the distance, it is $\frac{1}{9}$ as great, etc.

The falling apple is 6370 km from the center of the earth, and the moon, at an average distance of 384,000 km, is farther by factor of 60.3. We should expect that the acceleration of the moon would be smaller than that of the apple by a factor of 60.3^2, or 3636. The fraction, 980/3636, equals 0.270 cm/sec^2, a figure which checks very closely with the 0.273 cm/sec^2 we obtained from purely geometrical considerations of the moon's orbit. (The agreement would have been exact if we had not avoided mathematical difficulties by neglecting a minor complication. In figuring the centripetal acceleration of the moon, we tacitly assumed the earth to be stationary while the moon circled around it; actually, both the moon and the earth revolve around their center of gravity, located about 3000 miles from the earth's center.)

Newton also tested his ideas on the orbits of the planets around the sun and finally proposed his famous *law of universal gravitation*: ***every particle of matter in the universe attracts every other particle with a force proportional to the product of the masses of the particles, and inversely proportional to the square of the distance between them.***

In shorter algebraic form,

$$F = \frac{Gm_1m_2}{d^2}.$$

In Newton's day, it was not possible to evaluate the proportionality constant G, because the force of attraction between any masses that can be handled in the laboratory is very small indeed. A century after Newton, Henry Cavendish used a special type of balance to determine the value of G by measuring the force of attraction between known masses. Later more accurate work has given us the value

$$G = 6.67 \times 10^{-8}$$

in CGS units. This means that two 1-gm masses 1 cm apart will attract each other with a force of 6.67×10^{-8} dyne.

Once G is known, the mass of the earth can readily be determined. The force of attraction between the earth and a 1-gm mass is known to

be 980 dynes, and the 1-gm mass is separated from the earth's center by a distance equal to the earth's radius, which is 6.37×10^8 cm. Substituting these values in the equation, we get:

$$980 = \frac{6.67 \times 10^{-8} \times 1 \times m_e}{(6.37 \times 10^8)^2}$$

$$m_e = 5.97 \times 10^{27} \text{ gm}.$$

6-2 Gravitational Potential

It is easy to visualize the kinetic energy of the moon or of a satellite hurtling at high speed about the earth. We must not overlook, though, the fact that these bodies also have potential energy as well as kinetic energy. In our previous dealings with gravitational potential energy, it was necessary only to multiply the weight of a body by the height it was lifted above some chosen reference plane: the distances involved have been small enough for us to consider the pull of gravity to be constant. However, when we deal with distances of thousands of miles above the earth's surface, the pull of gravity is materially reduced, and this reduction must be taken into account.

Let us take the general case of two masses, M and m, and derive an expression for the work we must do against their gravitational attraction if we want to separate them. Figure 6-1 shows M and m; we shall compute the work needed to move m from point a to point f by dividing the whole distance into small steps and figuring the work for each small step separately.

The gravitational attractive force when m is at a will be GMm/r_a^2, where G is Cavendish's gravitational constant. At b, the force is GMm/r_b^2. If we take the *average* of the force at a and the force at b and multiply the average by the distance ab, we should get the work required to move m

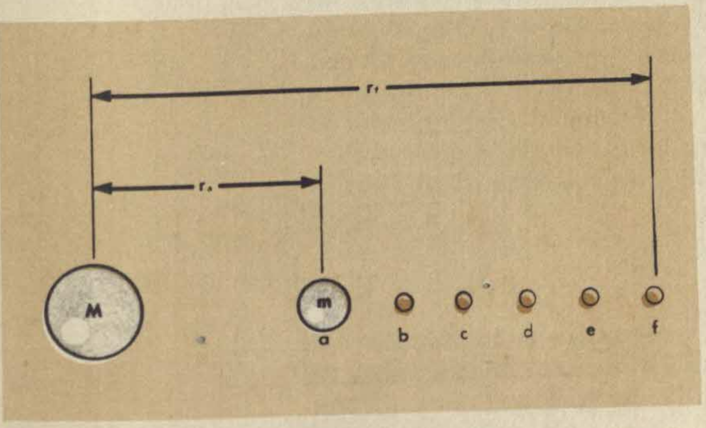


FIG. 6-1 The work done in moving apart masses that gravitationally attract each other.

from a to b. There are several ways to take an average, the most common being to get the *arithmetic mean* by adding the two forces together and dividing by two. However, using this simple arithmetic mean would lead into some complicated mathematical expressions that would involve great difficulties. Instead, let us average them by taking the *geometrical mean*; to do this we multiply the forces together and take the square root of the product. (Fortunately for us, this not only works out nicely, but calculus would show us that it is also the correct method to use in this case.)

By this method, we find that the average force between points a and b, which we can call F_{ab}, is

$$F_{ab} = \sqrt{\frac{GMm}{r_a^2} \times \frac{GMm}{r_b^2}} = \frac{GMm}{r_a r_b}.$$

The work done in moving m from a to b will be the force times the distance, which is

$$\begin{aligned} W_{ab} &= F_{ab}(r_b - r_a) \\ &= \frac{GMm}{r_a r_b}(r_b - r_a) \\ &= GMm\left(\frac{1}{r_a} - \frac{1}{r_b}\right). \end{aligned}$$

Similarly, the work done in moving m from b to c is

$$W_{bc} = GMm\left(\frac{1}{r_b} - \frac{1}{r_c}\right)$$

and so on until we get to point f.

We can now get the total work by adding all these pieces together. (Since each expression is multiplied by the factor GMm, let us lay GMm aside, and try not to forget to put it back when we are finished.) This gives

$$\begin{aligned} &\frac{1}{r_a} - \frac{1}{r_b} \\ +\ &\frac{1}{r_b} - \frac{1}{r_c} \\ +\ &\frac{1}{r_c} - \frac{1}{r_d} \\ +\ &\frac{1}{r_d} - \frac{1}{r_e} \\ +\ &\frac{1}{r_e} - \frac{1}{r_f} \\ \hline =\ &\frac{1}{r_a} - \frac{1}{r_f}. \end{aligned}$$

Or, putting our laid-aside factor back,

$$W_{af} = GMm\left(\frac{1}{r_a} - \frac{1}{r_f}\right).$$

With the above expression, we are able to compute the work needed to move m from any distance r_a from M, to any other distance r_f. The expression gives a somewhat surprising answer to the question, "how much work would be needed to move m to an infinite distance away?" The unthinking snap-judgment answer—"an infinite amount of work, of course."—sounds reasonable but is actually very wrong. Because of the inverse square law of gravitation, as we go out farther and farther from any attracting body, the force gets small more rapidly than the distance gets large. As a result, the question has a definite, finite answer. As m is moved farther and farther away, r_f becomes increasingly large, and $1/r_f$ becomes smaller and smaller. At $r_f = \infty$ (∞ is the mathematician's symbol for infinity), $1/r_f$ becomes $1/\infty$, or zero. The amount of work needed to move m from point a to an infinite distance away thus becomes simply

$$W_{a\infty} = \frac{GMm}{r_a}.$$

6-3 Escape Velocity

A great deal has been written in the newspapers and in popular articles about *escape velocity*. It is an important idea these days when man is commencing to think seriously of the time in the not-too-far-distant future when he may be able to leave the earth for new explorations in space. Escape velocity from the earth is given as about 7 mi/sec and is usually defined to be the speed a projectile must have in order to "break the bonds of gravity" or to "escape the earth's gravitational pull." Since the earth's gravitational pull extends to infinity (however weak it may be at great distances), escape velocity is apparently the velocity a projectile must have at the earth's surface in order for it to be projected an infinite distance into space.

From the previous section, we know that the work required to move a body from the earth's surface (r_e from the center of the earth) to infinity is

$$W = \frac{GMm}{r_e}.$$

If a projectile is given this amount of kinetic energy, it will have reached an infinite distance before its KE has been converted into PE; in other words, it will have *escape velocity*. With any smaller speed it will ultimately stop and fall back.

As in earlier problems, escape velocity can be readily computed by setting PE = KE:

$$\frac{GMm}{r_e} = \frac{1}{2}mv^2$$

$$v^2 = \frac{2GM}{r_e}.$$

The letter M, of course, is the mass of the earth, which is 5.98×10^{27} gm. The fact that m cancels out of the formula shows that escape velocity is the same for bodies of all masses; whether they are molecules of the upper atmosphere or spaceships makes no difference. The radius of the earth, 6.37×10^8 cm, is represented by r_e, and Cavendish's gravitational constant G is 6.67×10^{-8} in the CGS units we have used for M and r_e. By substituting these values in our equation we get

$$v^2 = \frac{2 \times 6.67 \times 10^{-8} \times 5.98 \times 10^{27}}{6.37 \times 10^8} = 12.52 \times 10^{11}$$

and

$$v = 1.12 \times 10^6 \text{ cm/sec}$$
$$= 11.2 \text{ km/sec}$$

or

$$= 11.2 \times 0.6214 = 6.96 \text{ mi/sec}.$$

The velocity needed to escape from another planet will in general be different from the 7 mi/sec that we have figured for escape from the surface of the earth. It must be computed separately for each planet and will depend on the planet's radius and mass.

The speed of a satellite in its orbit, whether it be natural or man-made, must obviously be less than escape velocity or it would fly away into infinite space. There is a definite relation between the orbital speed of a satellite and the speed it would need to have in order to escape from that orbit and travel away into infinity. We know from our earlier discussion of the moon and centripetal acceleration that the centripetal force holding a satellite in its orbit *is* the gravitational attraction of the planet it circles around. For an earth satellite of mass m and orbital radius r, equating the gravitational pull with the centripetal force gives

$$\frac{GMm}{r^2} = \frac{mv^2}{r}$$

from which

$$v_o^2 = \frac{GM}{r}.$$

We already have figured that the escape velocity from a distance r from the earth's center is

$$v_{\text{esc}}^2 = \frac{2GM}{r}.$$

A comparison of the above equation with the one preceding it shows that

$$v_{\text{esc}}^2 = 2v_{\text{orb}}^2$$

or

$$v_{\text{esc}} = \sqrt{2}\, v_{\text{orb}}.$$

That is, the velocity needed to escape from any distance r from the earth's center is just 1.414 times the speed of a satellite orbiting at that same distance.

6-4 Orbits

So far, our calculating has been done as though the orbits we considered were perfectly circular. For many purposes this is a good enough approximation, although scientists have known for more than 300 years that the orbits of planets and satellites are ellipses. The discovery of the ellipticity of orbits was made by the German mathematician and astronomer Johannes Kepler early in the seventeenth century. He set forth his discoveries in the form of the three following laws:

I. ***Each planet revolves around the sun in an elliptical orbit, with the sun at one focus of the ellipse.***

II. ***The speed of a planet in its orbit varies in such a way that the radius connecting the planet and the sun sweeps over equal areas in equal times.***

III. ***The squares of the periods of any two planets are in the same ratio as the cubes of their average distance distances from the sun.***

Mathematically, Kepler's third law can be stated as

$$\frac{T_1^2}{T_2^2} = \frac{d_1^3}{d_2^3}.$$

(The period of a planet is the time the planet takes to make one complete revolution around the sun.)

It can be shown (again, by the use of calculus!) that if the sun attracts a planet with Newton's inverse square law of gravitational force, the planet's orbit must necessarily be an ellipse, with the sun at one focus. The orbit may be practically circular, as in the case of the planet Venus, or have a very elongated shape, like the orbits of many comets, but it must always be an ellipse.

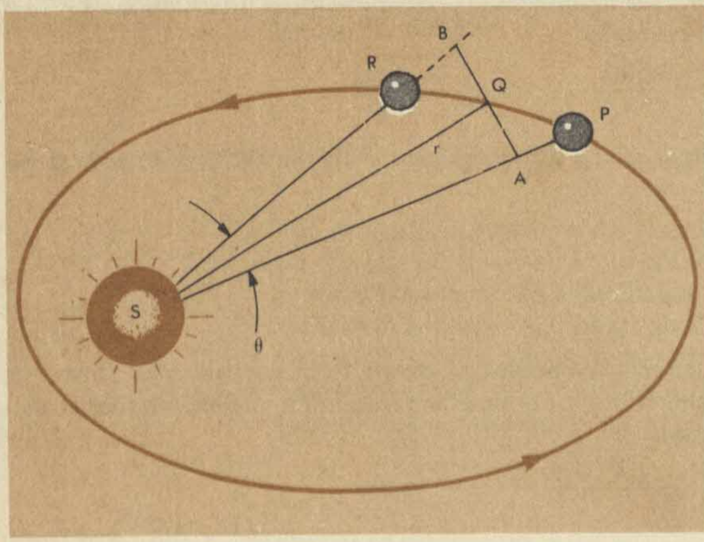


FIG. 6-2 Kepler's second law.

Kepler's second law necessarily follows from the principle of conservation of angular momentum. Figure 6-2 shows a planet which has moved along its orbit from P to R in some definite short time interval t. The radius connecting the planet to the sun has swept over the area PSR during time t. We can compute this area by drawing the radius SQ to the midpoint between P and R. Then AB, perpendicular to SQ, forms another triangle ASB, with an area equal to the area of PSR. The area of ASB is $\frac{1}{2}r \times AB$. But AB equals $r\theta$, where θ is the angle through which the planet has traveled during time t; and $\theta = \omega t$. All these substitutions put together give us

$$\begin{aligned} \text{area} &= PSR = ASB = \tfrac{1}{2} \times r \times AB \\ &= \tfrac{1}{2} \times r \times r\theta \\ &= \tfrac{1}{2}r^2\,\omega t. \end{aligned}$$

A slight rearrangement gives

$$\text{area} = t \times \frac{r^2\omega}{2}.$$

Let us now multiply both numerator and denominator of the fraction by m, the mass of the planet:

$$\begin{aligned} \text{area} &= t \times \frac{mr^2\omega}{2m} \\ &= t \times \frac{I\omega}{2m}. \end{aligned}$$

The expression $I\omega$ is the angular momentum of the planet; no matter how strongly or weakly the sun and the planet attract each other along the direction of the line connecting them, the angular momentum cannot be changed. The principle of conservation of angular momentum tells us that, unless some *outside* force acts on the system, its angular momentum must remain constant. The radius r will increase and decrease as the planet's distance from the sun changes, but the angular speed ω will always change just enough to keep the product $mr^2\omega$ everywhere the same.

Since $I\omega$ is always the same and the mass of the planet m does not change, we can tell from the equation that during equal time intervals t, the area swept over by the radius must be the same in any part of the planet's orbit.

Kepler's third law follows much more directly from the laws of mechanics. Let us designate by a subscript 1 a satellite revolving about a parent body of mass M. We have already shown that

$$v_1^2 = \frac{GM}{r_1}.$$

The satellite's average velocity also equals the circumference of its orbit divided by the period T_1:

$$v_1 = \frac{2\pi r_1}{T_1}$$

$$v_1^2 = \frac{4\pi^2 r_1^2}{T_1^2}.$$

We can now set these two different expressions for v_1^2 equal to each other and get

$$\frac{GM}{r_1} = \frac{4\pi^2 r_1^2}{T_1^2}$$

or

$$\frac{T_1^2}{r_1^3} = \frac{4\pi^2}{GM}.$$

For another satellite of the same parent body, we would also get

$$\frac{T_2^2}{r_1^3} = \frac{4\pi^2}{GM},$$

and thus it follows that

$$\frac{T_1^2}{r_1^3} = \frac{T_2^2}{r_2^3},$$

or, if we prefer to rearrange the equation a bit,

$$\frac{T_1^2}{T_2^2} = \frac{r_1^3}{r_2^3}.$$

By the use of Kepler's third law, together with what we know about the moon, we can readily determine the period of any other satellite of the earth. For example, consider an artificial satellite whose perigee (closest point to the earth) is 3000 miles above the earth's surface, and whose apogee (farthest point from the earth) is 4200 miles above the surface. Its *average* distance is 3600 miles, which is $3600 + 3960 = 7560$ miles from the earth's center. We know that the moon is about 239,000 miles from the center of the earth and has a period of 27.3 days. Using Kepler's third law, we get

$$\frac{T_{\text{sat}}^2}{(27.3)^2} = \frac{(7560)^3}{(239{,}000)^3}$$

or

$$\begin{aligned} T_{\text{sat}}^2 &= \frac{(27.3)^2 \times (7560)^3}{(239{,}000)^3} \\ &= \frac{(2.73)^2 \times 10^2 \times (7.56)^3 \times 10^9}{(2.39)^3 \times 10^{15}} \\ &= 236 \times 10^{-4}. \\ T_{\text{sat}} &= 0.154 \text{ day} \\ &= 3.69 \text{ hr.} \end{aligned}$$

Questions

(6-1)

1. A young man (mass 80 kg) stands with his center of gravity 0.5 m from that of a girl whose mass is 50 kg. What is the *gravitational* attraction between them?

2. A ship of 2×10^4 metric tons is moored 50 m (center to center) from a ship of 4×10^4 metric tons. With what force do they gravitationally attract each other?

3. If a body were released from a spaceship at a distance of 1.92×10^5 km from the earth's center (this is half the distance to the moon), with what acceleration would it fall toward the earth?

4. At an elevation of 6370 km above the earth's surface (this is equal to the radius of the earth), with what acceleration would a body fall toward the earth?

5. What are the dimensions of G (the constant in the gravitational formula) in CGS units?

6. What is the numerical value of G in MKS units?

(6-2)

7. Take two numbers relatively far apart, say 4 and 16. How does their arithmetical mean compare with their geometrical mean?

8. Take two numbers relatively close together, say 24 and 25. How does their arithmetical mean compare with their geometrical mean?

9. (a) How much work would be required to lift a load of 10^4 kg from the earth's surface to an altitude above the surface equal to the earth's radius? (b) How does this compare with the erroneous answer we would get if we figured

the earth's gravitational pull to remain constant at the same value it has at the surface?

10. (a) How much work would be required to raise a mass of 1000 kg from the earth's surface to an altitude above the surface equal to the earth's diameter? (b) How does this compare with the erroneous answer we would get if we figured the earth's gravitational pull to remain constant at its surface value?

11. (a) How much work would be needed to remove a gram from the earth's surface to infinity? (b) If we decide to call the PE zero when at infinity, what is the PE of the gram when on the earth's surface?

12. (a) How much work would be needed to remove a mass of 2 kg from the earth's surface to infinity? (b) If we decide to call the PE zero when at infinity, what is the PE of the 2-kg mass when on the earth's surface?

(6-3)

13. Escape velocity from the earth's surface is 11.2 km/sec. (a) What would the escape velocity be if the earth's mass were doubled, its size being unchanged? (b) What would the escape velocity be if the earth's radius were doubled, its mass being unchanged?

14. Escape velocity from the earth's surface is 6.96 mi/sec. (a) What would the escape velocity be if the earth's mass were multiplied by 9, its size being unchanged? (b) What would the escape velocity be if the earth's radius were multiplied by 4, its mass being unchanged?

15. What is the escape velocity from the surface of Planet X, which has 16 times the earth's mass and whose diameter is 9 times that of the earth?

16. The mass of Mars is 0.108 times the earth's mass, and it has a radius 0.532 times the radius of the earth. What is the escape velocity from Mars?

17. What would be the speed of a satellite orbiting around the planet of Question 15, at a small distance above its surface?

18. What would be the speed of a satellite orbiting around Mars at a small distance above its surface? (See Question 16.)

19. What is the speed of a satellite whose circular orbit is 1000 miles above the earth's surface? (Assume this distance to be $\frac{1}{4}$ radius.)

20. What is the speed of a satellite in a circular orbit at an elevation above the surface equal to the earth's diameter?

(6-4)

21. The planet Mercury has a very elliptical orbit. At perihelion (closest to the sun) it is 28.6×10^6 mi from the sun and has a speed of 35 mi/sec. At aphelion (farthest from the sun) it is 43.4×10^6 mi from the sun. What is its speed at aphelion?

22. From March 21 to September 22 the earth moves 180° in its orbit around the sun. Count the number of days required to cover each half of the orbit. From this, is the earth closest to the sun in winter or in summer?

23. Mars has two moons: Deimos, the larger, orbits at a mean distance of 6.9 Martian radii from the center of Mars, and its period is about 30 hours. Phobos, the smaller moon, has a period of approximately 7.6 hours. How far (in Martian radii) is Phobos from the center of Mars?

24. If the earth had an artificial satellite with an orbit of 119,500 mi radius

(half the radius of the moon's orbit), what would be its period of revolution around the earth?

25. What would be the radius of the orbit of a Martian satellite with a period of 20 hr? (Give answer in units of Martian radii. See Question 23.)

26. The earth now has communications satellites for relaying radio and TV signals from one continent to another. Viewed from the earth these satellites appear to hang nearly motionless in the sky directly over some point on the Equator. (a) What is the period in which they revolve about the earth's center? (b) What is the approximate radius of their orbits? (c) How far are they above the earth's surface?

chapter / seven

Elastic Vibrations

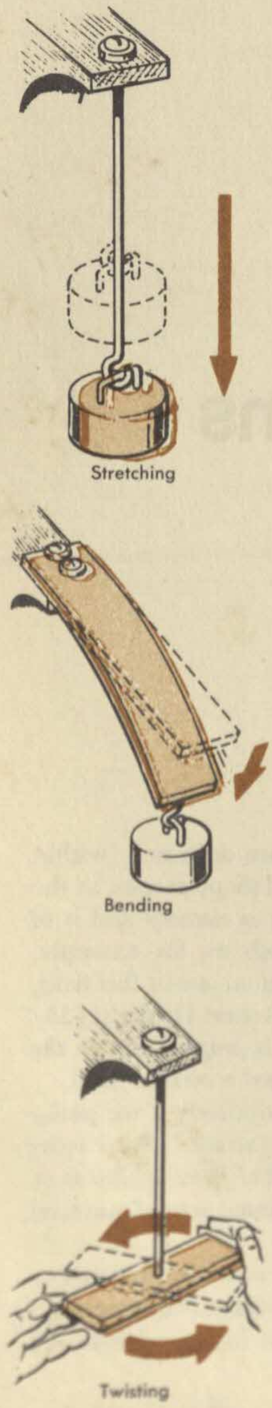
7-1 Young's Modulus

Solid bodies tend to maintain their shape and when deformed (within limits) by an external force, return to their original shape as soon as the force is removed. This property of solids is known as *elasticity* and is of great importance in many practical devices, such as, for example, those based on various types of springs. A fundamental law in this field, *Hooke's law of elasticity*, named after its discoverer Robert Hooke (1635–1703), states that ***the deformation of a solid body is proportional to the force acting on it***, provided the force does not exceed a certain limit.

It will be easier to deal with Hooke's law quantitatively if we pause to define a pair of technical terms, "stress" and "strain." *Stress refers to the internal forces created within a material as a result of forces applied to it.* It is F/A, the applied force divided by the cross-sectional area of material that resists the force.

Imagine a weight of F dynes hung from the end of a wire that is 1 m long and that has a cross-sectional area of 2 mm^2. Since F is supported by 2 mm^2 of metal all along the wire, the stress in the wire is $F/2$

FIG. 7-1 Three types of elastic deformation.



dynes/mm^2. If the wire were thicker, with a cross-sectional area of 4 mm^2, there would be more material to support F, and the stress would be reduced to $F/4$ dynes/mm^2.

Strain is a measure of how much a body is deformed by a stress. The principal types of deformation are stretching (or compressing), bending, and twisting, as shown in Fig. 7-1. Let us look first at the simplest of these—the lengthening of a stressed wire. The strain of the wire is the fractional increase in length caused by the stress—that is, it is the elongation divided by the length. Mathematically, this is $\Delta l/l$, and since it is a length divided by a length, strain is a pure number without dimensions.

If we hang a weight on a piece of wire 1 m long, the wire will stretch a certain amount. The same weight hung on a piece of the same wire 2 m long will cause it to stretch twice as much, since each meter will elongate exactly the same as it did before. In both cases the strain $\Delta l/l$ is the same, and we can see that strain has the useful property of being independent of the length of the wire involved.

We can restate Hooke's law in a more useful way by saying, "stress is proportional to strain," or

$$\frac{\text{stress}}{\text{strain}} = \text{a constant.}$$

This is true within certain limits. If the wire is stressed too much, it will no longer return to its original length when the stress is removed, and we say that we have exceeded the *elastic limit* of the material. Within the elastic limit, however, we can expand the equation above into

$$\frac{F/A}{\Delta l/l} = Y.$$

The proportionality constant Y is called the *Young's modulus* of the material, after a great English engineer, scientist, and philosopher of the early nineteenth century. We shall see his name again later in connection with optics. Here is Young's modulus for a few common materials:

Steel	2×10^{12} dynes/cm^2	or	3×10^7 lb/in.^{2}
Aluminum	7×10^{11}		1×10^7
Copper	1×10^{12}		1.5×10^7
Wood	1×10^{11}		1.5×10^6

Since strain is dimensionless, it is apparent that Young's modulus must have the same dimensions as the stress—that is, force per unit area.

Let us see how this applies to a copper wire 0.4 mm in diameter and 3.0 m long. How much will this wire elongate if we use it to suspend a weight of 5 kg? First, we note that Young's modulus for copper is given in dynes/cm^2, so we must figure the stress on the wire in the same units.

Since a gram weighs 980 dynes, the force is $5\text{ kg} \times 1000 \times 980 = 4.90 \times 10^6$ dynes. The circular cross-sectional area of the wire that resists this pull is $\pi r^2 = 3.14 \times (0.02)^2 = 1.26 \times 10^{-3}\text{ cm}^2$. From these values of force and area we can get the stress, which is $4.90 \times 10^6/1.26 \times 10^{-3} = 3.89 \times 10^9\text{ dynes/cm}^2$.

Since Young's modulus was defined as $Y = \text{stress/strain}$, it is apparent that strain = stress/Y, and we have

$$\text{strain} = \frac{3.89 \times 10^9}{1 \times 10^{12}} = 3.89 \times 10^{-3}.$$

This figure can be interpreted to mean that any piece of copper under a stress of $3.89 \times 10^9\text{ dynes/cm}^2$ will elongate 3.89×10^{-3} of its length (or, if it is compressed, will shorten by the same fraction). Thus the elongation of the wire will be

$$\begin{aligned}\Delta l &= \text{strain} \times l \\ &= 3.89 \times 10^{-3} \times 3.0 \\ &= 1.17 \times 10^{-2}\text{ m}\end{aligned}$$

or

$$= 1.17\text{ cm}, \quad \text{or} \quad 11.7\text{ mm}.$$

The elongation or the shortening of a coiled spring cannot be so easily figured from a knowledge of just its material and dimensions. As the spring is stretched, the principal stress on the wire of which it is made is twisting, and to this twisting is added some bending, as well as a small amount of straight tensile pull. It is too complicated a problem for us to go into, but we should note that within its elastic limit, a spring will nevertheless obey Hooke's law.

7-2 Simple Harmonic Motion

If we suspend a chunk of iron or lead from a rubber band or steel spring attached to the ceiling and by a gentle push set it bobbing up and down, we will find it moves with a very definite *period*. The period T of a regularly repeating motion such as this is the time needed to make a complete up-and-down cycle, and is obviously the reciprocal of the frequency f, which is the number of cycles or complete vibrations the body makes in a unit of time:

$$T(\text{sec/vibr}) = \frac{1}{f(\text{vibr/sec})}\cdot$$

If the spring or rubber band obeys Hooke's law, the motion of the chunk of iron will be an example of *simple harmonic motion* (SHM); so also will be the oscillation of a diving board, the vibration of a plucked violin string, or the swinging of a pendulum. Besides describing such back-and-forth movements, SHM lies behind all wave motion, and it

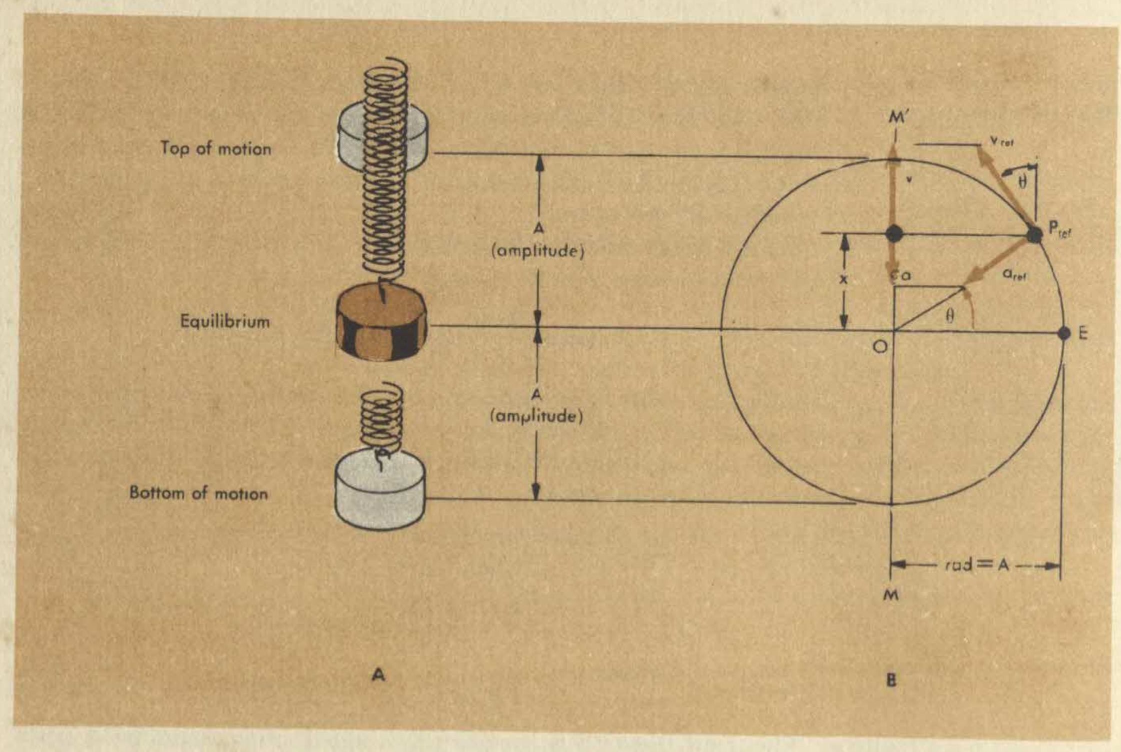


FIG. 7–2 Simple harmonic motion and the reference circle.

will be worth our while to look at this special kind of motion more closely.

Figure 7-2A shows a weight suspended by a spring; with the weight hanging motionless, it is in the position marked "equilibrium." Here the upward pull of the spring just equals the pull of gravity on the weight. Now pull the weight down to "bottom of motion" and let it go; it will oscillate between "bottom of motion" and "top of motion" in SHM. The total range of travel of the weight will be $2A$. (A, the *amplitude*, is the distance from the equilibrium point to either extreme of motion.)

Let us temporarily forget the bobbing weight and turn to a piece of geometrical fiction shown in Fig. 7-2B. On a line MM' parallel to the motion of the weight mark a point O level with the equilibrium position. Now with O as center, draw a circle with a radius equal to A, the amplitude of the motion. Around this circle (called the *reference circle*), imagine that a tiny *reference particle* P_{ref} is rotating with a constant angular velocity ω. If we click an imaginary stopwatch and thus start counting time, at the instant when P_{ref} passes E, the angle θ will be equal to ωt. The velocity of P_{ref} is represented by the vector v_{ref}, with a magnitude $A\omega$. We know that P_{ref} must have a centripetal acceleration, represented by the vector a_{ref}, which is pointed toward the center of the circle and has a magnitude $A\omega^2$.

We can now turn from P_{ref} to another particle P, whose motion will duplicate that of the bobbing weight. Particle P oscillates back and forth along MM' in such a way that its position (distance x from the center), its velocity v, and its acceleration a are *projections* of P_{ref}, v_{ref}, and a_{ref} on the line MM'. (A point, say P_{ref}, is projected on MM' by drawing a line from P_{ref} perpendicular to MM'. Where this perpendicular hits MM', at P, is the projection. Vectors are projected by projecting their end points, as shown in Fig. 7-2B.) Another way of visualizing this is to think of x, v, and a as being the components of OP_{ref}, v_{ref}, and a_{ref} in the direction of the line MM'. We see that

$$x = OP_{ref} \sin\theta = A \sin\omega t$$

$$v = v_{ref} \cos\theta = \omega A \cos\omega t$$

$$a = a_{ref} \sin\theta = -\omega^2 A \sin\omega t = -\omega^2 x.$$

The minus sign is necessary in the last equation because the direction of the acceleration is always opposite to the displacement of x from the central equilibrium point.

A moment's reflection (and possibly, also, another look at Fig. 2-9) will show that the magnitude of the sine can vary from 0 (at 0°) to 1 (at 90° or $\pi/2$ radians); the magnitude of the cosine varies from 0 (at 90° or $\pi/2$ radians) to 1 (at 0°). From this observation and the equations above, we can see that the maximum value of x is A, which occurs at $\omega t\ (= \theta) = 90°$, and $x = 0$ at $\theta = 0$. The maximum speed of P is ωA, at the equilibrium point, when $\theta = 0$, and is zero at maximum displacement ($\theta = 90°$). The acceleration a varies from zero at the midpoint to a maximum of $\omega^2 A$ at maximum displacement.

If P is to duplicate the motion of the suspended weight, θ must go through 2π radians while the weight goes through one cycle—that is, ω (in rad/sec) equals $2\pi f$ (in cycles, or vibrations, per sec). It will be more useful if the equations above are rewritten to be in terms of f rather than ω:

$$x = A \sin 2\pi f t$$

$$v = 2\pi f A \cos 2\pi f t$$

$$a = -4\pi^2 f^2 A \sin 2\pi f t = -4\pi^2 f^2 x.$$

The last of these equations gives us an opportunity to connect our geometry with the actual physical characteristics of the spring and the weight. From $F = ma$, we get (omitting the minus sign)

$$F = 4\pi^2 f^2 x m$$

or

$$\frac{F}{x} = 4\pi^2 f^2 m.$$

Since the period T (sec/vibr) $= 1/f$ (vibr/sec), this equation can be rewritten as

$$\frac{F}{x} = \frac{4\pi^2 m}{T^2}$$

from which

$$T^2 = \frac{4\pi^2 m}{F/x}$$

and

$$T = 2\pi\sqrt{\frac{m}{F/x}}.$$

The term F/x is the force per unit stretch of the spring, which tends to bring the mass m back to its equilibrium position when it is moved away from this position in either direction. If the spring obeys Hooke's law (and in general it will if it is not stretched too far), this ratio F/x will be the same for all amounts of stretch and is often called the *force constant* of the spring, or more simply the *spring constant.*

This last equation above is remarkable for the things it does *not* contain. The amplitude A is missing, and so is g, the acceleration of gravity. The equation thus tells us that the frequency with which the weight bobs up and down depends only on the mass of the weight and on the force constant of the spring. The frequency is the same for motion of large amplitude or small and would be the same on the moon as it is on the earth.

As an example of SHM, let us take a bird watcher who sees a large bird light on the end of a slender tree limb, which is thus started oscillating. The bird makes 6 complete up-and-down bobs in 4 sec. When the bird leaves, the watcher hangs a 1-kg weight on the limb and measures that the weight deflects the limb 12 cm. What was the mass of the bird? Since a 1000-gm weight deflects the limb 12 cm, the force constant of the limb is $1000 \times 980/12$, or 8.17×10^4 dynes/cm. The frequency f is 6 vibr/4 sec = 1.5 vibr/sec. Whipping out his notebook, the ornithologist writes

$$8.17 \times 10^4 = 4\pi^2(1.5)^2 m$$

from which

$$m = 920 \text{ gm}.$$

7-3 The Simple Pendulum

A simple pendulum theoretically consists of a massive particle of zero size, swinging at the end of a massless rod or string. A small, compact piece of heavy metal suspended from a thread comes close enough to these requirements for most practical purposes. Figure 7-3 is a diagram of a simple pendulum, made by suspending a small ball of mass m from

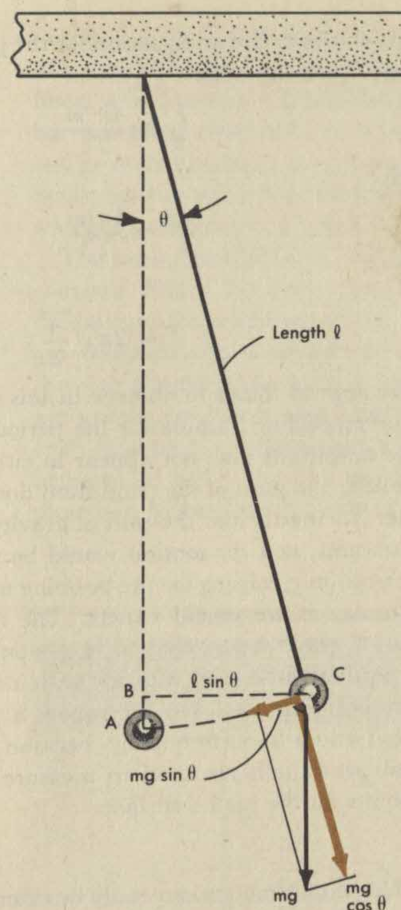


FIG. 7–3 The simple pendulum.

the end of a thread of length l. The diagram shows it pulled aside from the vertical by an angle θ. We can resolve the weight of the ball mg into two components: $mg \cos \theta$, which pulls directly against the thread and has no tendency to make the ball move; and a component $mg \sin \theta$ at right angles to the thread, which urges the ball toward equilibrium. The horizontal displacement of the ball from its equilibrium position is $l \sin \theta$. We can see that F/x for the ball is $mg \sin \theta / l \sin \theta = mg/l$, which is a constant, because m, g, and l remain unchanged in value. Thus the restoring force is proportional to the displacement of the ball, and it moves in SHM. (We have calculated the displacement x as $l \sin \theta$, which is horizontal; the restoring force F, however, is perpendicular to the string and is therefore not quite horizontal and not quite parallel to x. Hence the motion of the pendulum is not exactly SHM, but if θ is small, the difference is negligible.)

Let us use this to derive an equation for the period T of the pendulum. Since $f = 1/T$, we can write

$$\frac{F}{x} = \frac{4\pi^2 m}{T^2}$$

from which

$$\frac{mg}{l} = \frac{4\pi^2 m}{T^2}$$

and

$$T = 2\pi\sqrt{\frac{l}{g}}\,.$$

There are several things to observe in this formula, and to compare with the corresponding formula for the period of a mass bobbing on a spring. The amplitude does not appear in either equation. In the pendulum formula, the mass of the pendulum does not appear; if the mass were greater, its inertia and the pull of gravity would both increase by the same amount, and the motion would be unchanged. The gravitational acceleration g, missing for the bobbing mass, is present in the pendulum formula, as we should expect. The restoring force is *provided* by the pull of gravity; any change in g would have an effect on the pendulum equivalent to using a spring with a different force constant to suspend the bobbing mass. On the moon, a pendulum would have a longer period and a lower frequency, because the lunar g is smaller. In fact, special pendulums are used to measure small variations in g at different points on the earth's surface.

7-4 Rotational SHM

The idea of harmonic motion can easily be extended to include rotational motion. By analogy with translational motion, which we have already established,

$$\frac{F}{x} = 4\pi^2 f^2 m$$

becomes

$$\frac{\tau}{\theta} = 4\pi^2 f^2 I$$

or

$$\frac{\tau}{\theta} = \frac{4\pi^2 I}{T^2}\,.$$

in which τ is the applied torque; θ is the twist which τ causes, measured in radians; and I is the moment of inertia about the axis of the twist. We now have

$$T = 2\pi\sqrt{\frac{I\theta}{\tau}} \quad \text{or} \quad 2\pi\sqrt{\frac{I}{\tau/\theta}}\,.$$

Figure 7-4 shows such a device (called a *torsion pendulum*) which consists of a very light crossbar suspended at its center by a thin wire or quartz fiber. A ball of mass m is placed on each end of the crossbar. If the crossbar is rotated away from its equilibrium position, the twist of the fiber will provide a restoring torque proportional to the angular displacement; hence the bar will rotate back and forth in rotational SHM. Almost all watches and some clocks are made to operate with this sort of pendulum.

The torsion pendulum is also used in determining G, the gravitational constant. Figure 7-4 shows two large masses M that are placed close to the balls of the pendulum and in the same horizontal plane, so that the gravitational attractions between M and m will twist the pendulum through a small angle from its equilibrium position. This angle can be accurately measured, and when the stiffness of the supporting fiber (τ/θ) is known, the deflecting torque and hence the gravitational pull between M and m can be computed. The stiffness of the fiber is determined by removing the large masses M and measuring the period of

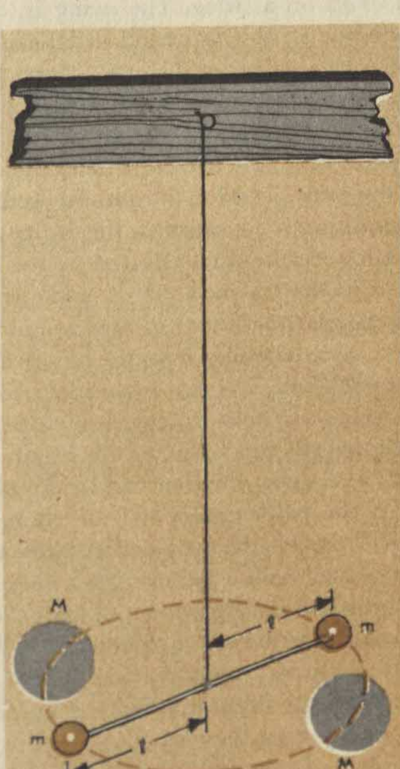


FIG. 7-4 A torsion pendulum, such as used to determine the gravitational constant G.

oscillation of the torsion pendulum when it is set to swinging by a gentle twist.

If, for example, the pendulum is made of two 10-gm balls on a rod 20 cm long and has a period of 20 min, we can compute the stiffness of the fiber in this way: the moment of inertia of the pendulum (ignoring the light rod) will be $2ml^2$, or $2 \times 10 \times 10^2 = 2000$ gm-cm². The period T is $20 \times 60 = 1200$ sec. A simple substitution gives

$$\frac{\tau}{\theta} = \frac{4\pi^2 \times 2000}{(1200)^2} = 0.0548 \text{ dyne-cm/rad}$$

In other words, a torque of 0.0548 dyne-cm would cause the fiber to twist through 1 radian, or about 57°.

7-5 Resonance

A very important notion in the study of all kinds of vibrations is that of *resonance*, which is *the specific response of a system which is able to oscillate or vibrate with a certain period, to an external force acting with the same period.* Consider a child on a swing. The swing is, of course, nothing but an ordinary pendulum, and its period is determined by the length of the ropes and is independent of the mass of the child. In order to put the swing in motion and to make it move with a larger and larger amplitude, the child must pull periodically on the ropes and stretch out his legs at the same time. But to be successful these muscular efforts must be made with the same period as the natural oscillation period of the swing, for if this condition is not satisfied, the swing will hardly move at all.

The situation can be demonstrated by the simple experiment shown in Fig. 7-5. A horizontal metal bar A is inserted loosely through the two holes in the supporting frame B, and a number of light wooden balls C_1, C_2, C_3, etc. are suspended from the bar on strings of different lengths. A heavier metal ball D is also suspended from the protruding end of bar A in such a way that its length of suspension can be changed. If we make this length equal, say, to the length of suspension of the third wooden ball and swing the iron ball in the plane perpendicular to the bar, some of the ball's energy will be transmitted through the small movements of the bar to the wooden balls hanging inside the frame. Although the other balls will show only a very slight tendency to take up this motion, the ball with the same suspension length (and, consequently, with the same oscillation period) will begin to swing more and more until its amplitude becomes even larger than that of the iron ball. By changing the suspension length of the iron ball, we can in turn set into motion the other balls hanging on the bar.

The principle of resonance is used in the construction of one type of an instrument called a *tachometer*, which is used for measuring the speed of rotation of various motors. It consists of a number of steel strips of

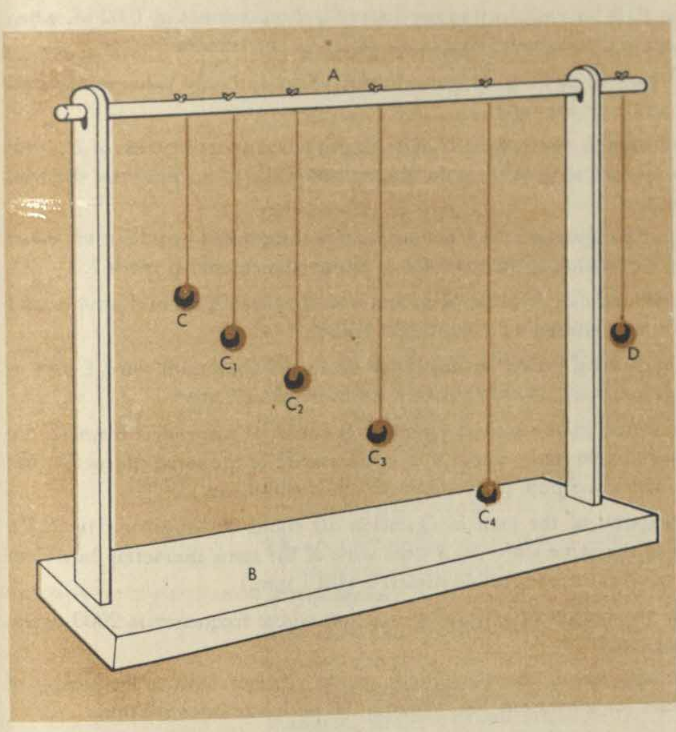


FIG. 7-5 The principle of resonance, as illustrated by a set of pendulums.

different lengths mounted on a common support. When put in contact with a running motor, the tachometer receives slight vibrations caused by the rotation of the motor's axis. The period of the motor's rotation will nearly coincide with the vibration period of one of the strips, and this particular strip will vibrate with an appreciable amplitude while all the other strips remain almost at rest. By reading the figures printed on a scale that runs along the row of strips, we can quickly find the number of revolutions per minute the motor is making.

Questions

(7-1)

1. A weight of 8000 lb is supported by a column 4 in. × 4 in. What is the stress in the column?

2. A load of 2000 kg is supported by a cylindrical column 10 cm in diameter. What is the stress in the column (a) in kg/cm^2? (b) in nt/m^2?

3. A wire 0.01 cm in diameter supports a load of 0.8 kg. What is the stress in the wire, in dynes/cm^2?

4. A load of 20 kg is suspended from a metal strip 1 cm × 0.02 cm. What is the stress in the strip, in dynes/cm^2?

5. A column 10 ft long supporting the floor of a shop shortens by 0.02 in. when a heavy machine is installed. What is the strain in the column?

6. A steel tape 100 m long elongates 2 mm when a tensile force is applied. What is the strain in the tape?

7. A piece of wire elongates by 10^{-3} of its length when a tensile stress of 2×10^9 dynes/cm^2 is applied to it. What is the Young's modulus of the material the wire is made from?

8. The column in Question 2 is 5 m long, and is compressed by 0.06 mm. What is the Young's modulus of the material of which the column is made?

9. How much load can be suspended from a steel wire 0.02 in. in diameter and 10 ft long, if it is to stretch no more than 0.1 in.?

10. How much load could be suspended from an aluminum wire 2 mm in diameter and 5 m long, if it is to stretch no more than 1 mm?

11. What fraction of the load of Question 9 could be suspended from (a) an aluminum wire of the same size? (b) a copper wire of the same diameter, but 20 ft long? (The maximum permissible stretch is still 0.1 in.)

12. What fraction of the load in Question 10 could be supported by (a) a copper wire of the same size? (b) a steel wire of the same diameter, but 15 m long? (The maximum permissible stretch is still 1 mm.)

(7-2)

13. What is the period of a vibratory motion whose frequency is 2000 cycles (or vibrations)/sec?

14. An electron in an electromagnetic wave vibrates with a frequency of 8×10^{14} cycles (or vibrations)/sec. What is the period of its oscillation?

15. The mass in SHM in Fig. 7-2 makes 10 cycles of motion in 4 sec. (a) What is its frequency? (b) What is its period? (c) What is the angular speed of the corresponding reference point (P_{ref}) in revolutions (or cycles) per sec? (d) What is the frequency (or angular speed) in radians/sec?

16. The mass in SHM in Fig. 7-2 makes 6 cycles in 30 sec. What is: (a) its frequency? (b) its period? (c) the angular speed of P_{ref} in revolutions (or cycles) per sec? (d) in radians/sec?

17. A weight suspended from a spring bobs up and down. At what point or points on its path is (a) its acceleration greatest? (b) its acceleration zero?

18. A weight suspended from a spring bobs up and down. At what point or points on its path is (a) its velocity greatest? (b) its velocity zero?

19. A reciprocating part on a machine moves in SHM with an amplitude of 5 cm and a frequency of 300 cycles/min. What is (a) its maximum velocity? (b) its maximum acceleration?

20. A bottle floating on the sea moves up and down in approximate SHM as the waves pass by; its motion has an amplitude of 8 inches and a frequency of 40 cycles/min. What is (a) its maximum velocity, in ft/sec? (b) its maximum acceleration?

21. A 500-gm mass is hung from the lower end of a coiled spring and causes the spring to elongate by 5 cm. If the mass is now set bobbing up and down, how many complete vibrations will it make in 1 min?

22. A mass of 4.8 lb is hung from the lower end of a coiled spring, and causes the spring to elongate by 3 in. If the mass is now set into vertical oscillation, what will its frequency be, in cycles/min?

23. Rework Question 21, with the same spring, but using a 1000-gm mass, and performing the experiment on the moon. (See Question 29.)

24. Rework Question 22, with the same spring, but using a 16-lb mass, and performing the experiment on the moon. (See Question 29.)

25. A weight is suspended by a string whose upper end is fastened to a machine part which moves up and down in SHM with an amplitude of 2 in. What is the maximum speed the machine can have if the string is not to become slack in any part of the cycle?

26. A flat horizontal platform moves up and down in SHM with an amplitude of 1 cm. A small object is placed on the platform. What is the maximum frequency the platform can have if the object is not to separate from it at any part of its motion?

(7-3)

27. What is the length of a simple pendulum which will "beat seconds," i.e., which will have a period of 2 sec?

28. A once-popular pendulum clock was the "80-beat" Seth Thomas clock, which made 40 complete oscillations/min. What is the length of the pendulum of this clock?

29. How long would it take the clock of Question 27 to record an hour if it were on the moon, where $g = 163$ cm/sec^2?

30. How long would it take the clock of Question 28 to record an hour on Planet Z, where g is 1470 cm/sec^2?

(7-4)

31. If, in Fig. 7-4 and the accompanying discussion in the text, the large balls each had a mass of 25 kg and were placed 15 cm (center-to-center distance) from the small balls, through what angle would the torsion pendulum twist?

32. The torsion pendulum which regulates the speed of a certain type of clock has a moment of inertia of 3000 gm-cm^2 and a period of 30 sec. How much torque would be needed to twist the suspending wire through an angle of 30°?

(7-5)

33. A piece of machinery weighing 3200 lb is mounted on four springs, each bearing one-fourth of the weight. The force constant of each of the springs is 2050 lb/in. Would it be advisable to operate this machine at a speed of 300 rev/ min? Explain.

34. A certain recording instrument whose mass is 12 kg is mounted on three springs to protect the instrument from vibration and shock. When the instrument is installed, each spring shortens by 15.5 mm. Would it be advisable to mount this recorder on a machine running at 240 rev/min? Explain.

chapter / eight

Waves

8-1 Wave Pulses

Probably the simplest wave we can imagine is a single disturbance, or pulse, started by a sideways jerk of the hand, which then travels down a long rope tied to a wall at its far end. It will be easy to experiment with this device, and by using a stopwatch to measure the time it takes for the wave to travel the length of the rope, you can easily compute the velocity of the wave. You will find that the tighter the rope is pulled, the faster the wave will travel. You will also find that if you substitute a rope of greater *linear density* (m_l)—that is, one that has a greater mass per unit length—the wave will travel more slowly than it did along the lighter rope pulled to the same tension. A little further investigation will show that the velocity of the wave is not affected by whether you jerk your hand quickly or slowly, or by whether you move it through a large or a small distance in starting the pulse, although the motion, of course, does influence the shape and size of the wave itself.

Study of Fig. 8-1 will make clear why wave velocity depends on the rope tension F and on the linear density m_l. Segment 1 has no choice;

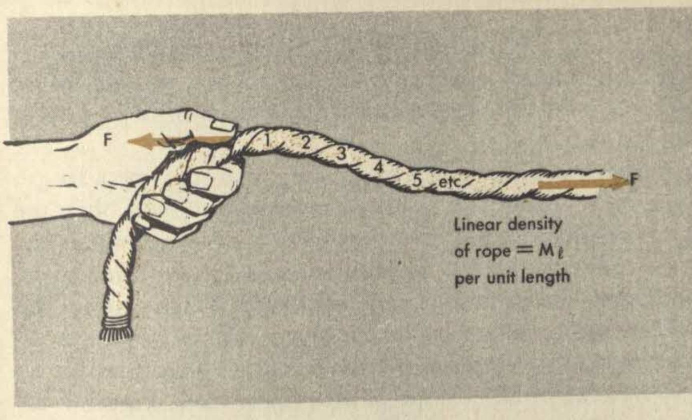


FIG. 8-1 Starting a transverse wave in a rope.

the hand forces it to move quickly upward and then back. In the drawing, segment 1 is at the top of the jerk and is about to start back. Segment 2 must follow along, and the greater the tension F, the more closely it will follow. Although segment 1 is about to start down, the kinetic energy of segment 2 will cause it to rise to the same distance as segment 1's maximum before it reverses its direction. Segment 3 follows segment 2, and so on, each segment duplicating the motion of the previous segment with a time delay which is less for a greater tension. Thus a greater tension results in a faster propagation of the wave. It is apparent, too, that if the mass of segment 2 is increased, the drag of segment 1 must operate for a longer time to bring segment 2 up to speed. This greater time delay in passing the motion along from segment to segment of the rope means that the wave travels more slowly in a rope of greater linear density. A mathematical derivation shows that velocity, tension, and linear density are related in the following way:

$$v = \sqrt{\frac{F}{m_1}}.$$

This formula applies to sideways waves of the sort we have been describing, traveling along a wire, a rubber tube, a long helical spring, a chain, or any other essentially one-dimensional wave carrier that is relatively flexible. It will *not* apply to stiff rods, planks, or similar objects in which the restoring force is due to the stiffness of the medium, rather than to a pure tensile stress such as we have considered in the example of the rope.

8-2 Wave Reflections

Up to now, we have neglected to think of what becomes of a wave, or pulse; we have merely started it and forgotten about it. In general, two things may happen to it. If the rope is long enough, the energy of the pulse may be gradually reduced by friction with the air and by the

internal friction of rope fibers rubbing against one another until the amplitude has been reduced to zero and the disturbance vanishes. More often, however, the pulse will be reflected, either entirely or in part, from a discontinuity in the rope.

We can make such a discontinuity by splicing onto our rope a length of rope with either a greater or a smaller linear density. As is often the case in physics, we shall do well to see what happens when these discontinuities are made as extreme as possible. This can be done easily by hanging a length of rope from a rigid support on a high ceiling; the upper end is fastened to something so massive it will not move at all when a wave reaches it, whereas at the lower end we have in effect an attachment to nothing at all—a rope of zero linear mass. Figure 8-2A shows a wave moving up such a rope. The wave, or pulse, was generated by the experimenter moving his hand first slowly to the left, then rapidly back again, so the wave begins with a gentle slope and ends with a steep

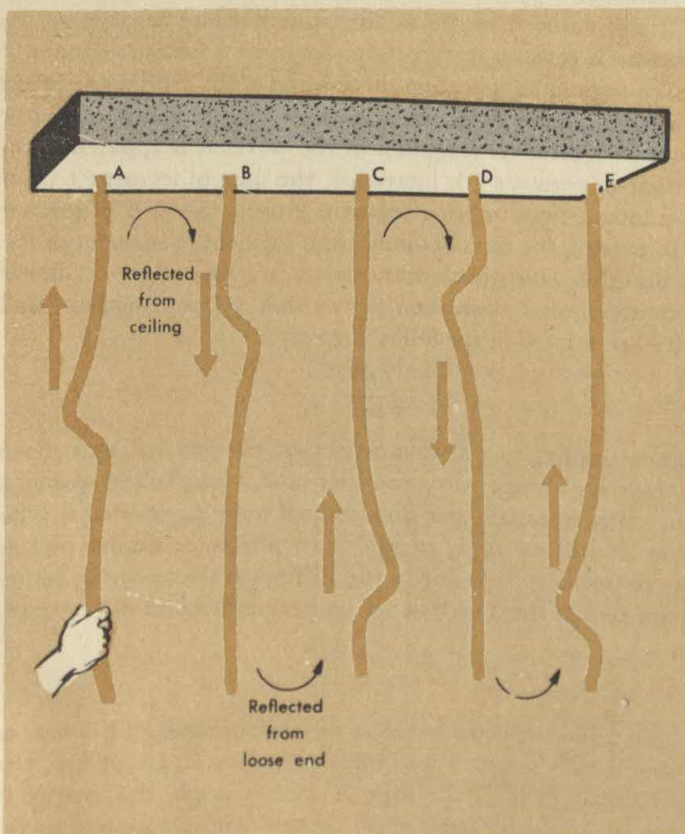


FIG. 8-2 Reflection of a pulse from the fixed end of a rope, and from the loose end.

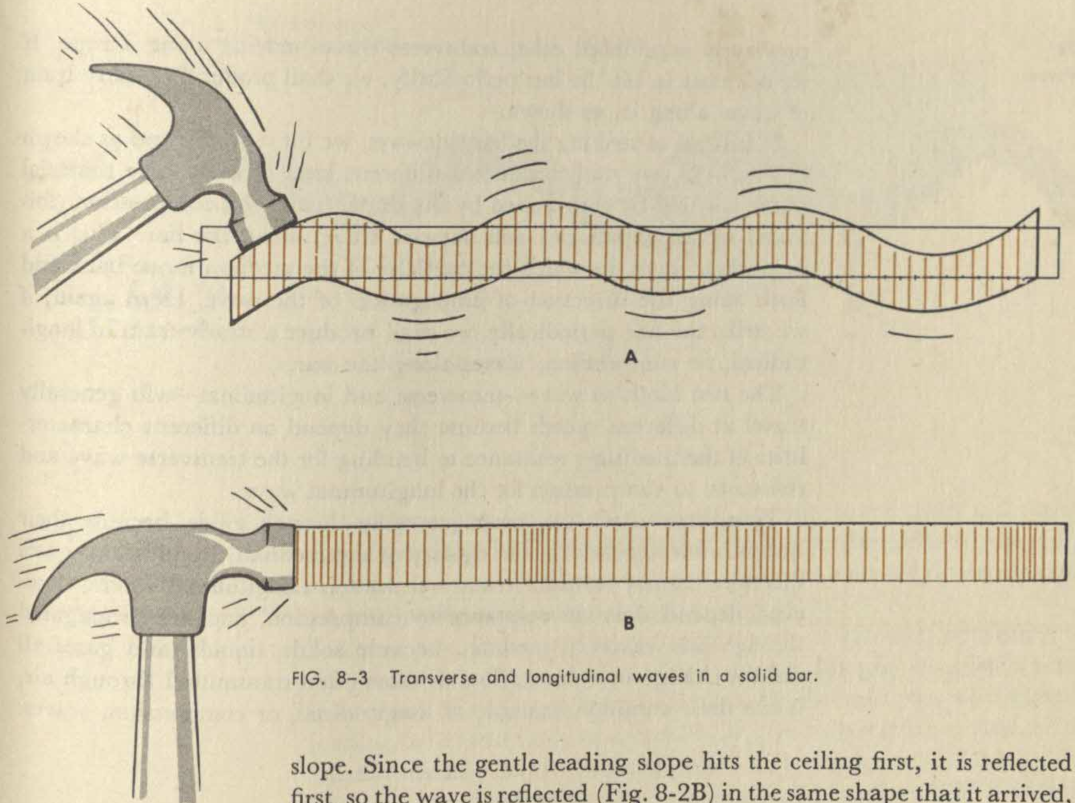


FIG. 8-3 Transverse and longitudinal waves in a solid bar.

slope. Since the gentle leading slope hits the ceiling first, it is reflected first, so the wave is reflected (Fig. 8-2B) in the same shape that it arrived, but on the opposite side of the rope. The direction of propagation of the wave has been reversed, and the direction in which the particles of the rope are displaced has also been reversed.

The reflection from the loose lower end is different; although the direction of the wave has, of course, been reversed (otherwise it would not be a reflection!), the pulse stays on the same side of the rope (Fig. 8-2C). After two more reflections (Fig. 8-2D and E), the wave appears again exactly as it began. (With any actual rope, however, the reflected pulse would be much smaller, owing to the dissipation of its energy by friction.)

8-3 Kinds of Waves

So far, we have considered only one special sort of wave—a pulse started by a sideways impulse, traveling down some one-dimensional medium. This is the simplest example of a *transverse* wave. As the disturbance passes along the rope or spring, the particles of the rope move at right angles—or transversely—to the direction of propagation of the wave. Figure 8-3A shows a metal bar being struck in a direction perpendicular to its length. The impact of the hammer will cause a bending-type deformation that will propagate along the bar just as we have

previously considered other transverse waves moving along a rope. If we continue to hit the bar periodically, we shall produce a steady train of waves along it, as shown.

If, instead of striking the bar sideways, we hit it on the end as shown in Fig. 8-3B, we shall produce a different kind of wave. The material of the bar will be compressed by the impact, and this compression (followed by an expansion) will likewise travel along the bar. This is a *longitudinal* wave, in which the particles of the medium move back and forth along the direction of propagation of the wave. Here again, if we strike the bar periodically, we shall produce a steady train of longitudinal, or compression, waves along the bar.

The two kinds of wave—transverse and longitudinal—will generally travel at different speeds because they depend on different characteristics of the medium: resistance to bending for the transverse wave and resistance to compression for the longitudinal wave.

Transverse waves can propagate only through solids, because their transmission depends on the rigidity of the medium. Liquids and gases therefore cannot transmit transverse waves. Longitudinal waves, however, depend only on resistance to compression, and are propagated through any material medium, because solids, liquids, and gases all resist a change in volume. Sound, most often transmitted through air, is the most common example of longitudinal, or compression, waves.

8-4 Periodic Wave Trains

Most wave phenomena of practical interest are concerned with long trains of equally spaced waves, rather than with the single pulses we have been observing in some detail. If the hand holding the rope in Fig. 8-4 moves up and down in simple harmonic motion, a snakelike series of waves will be sent traveling down the rope. A train of waves of this particular shape is called a *sine wave*, because the displacement of the rope is proportional to the sine of the SHM generating angle θ (see Fig. 7-2).

Such periodic trains of waves, whether they are sine waves or are of any other shape, have characteristics not possessed by simple pulses. If, say, the hand in Fig. 8-4 moves up and down f times per second, f complete waves per second will be generated. Wherever along the rope we choose to stand and count, we will find that f waves per second will pass by; this—*the number of complete waves that pass by any fixed point in a unit of time—is the frequency* of the waves. Alternatively, we could measure the time required for a single complete wave to pass. *This time interval, measured between the passage of one crest and the next crest—or between the passage of consecutive troughs—is the period* of the waves. It should be apparent that the frequency (f, in waves/sec) and the period (P, in sec/wave) are each the reciprocal of the other:

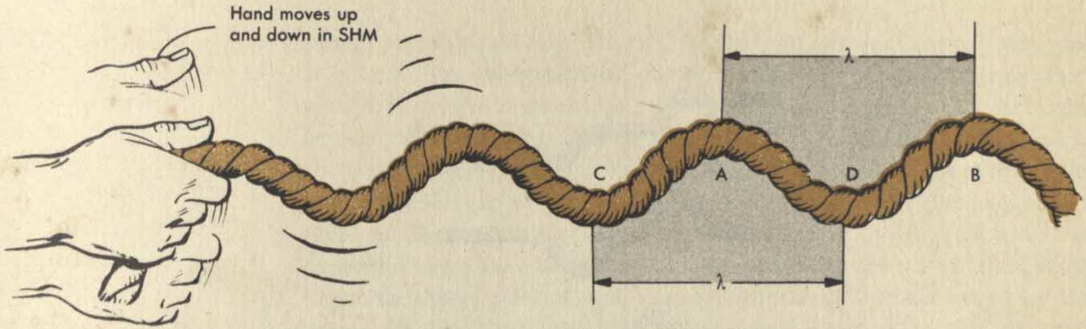


FIG. 8–4 Sine waves in a rope or wire.

$$f = \frac{1}{P} \quad \text{and} \quad P = \frac{1}{f}.$$

Another characteristic of a regular periodic wave train is its *wavelength*, generally indicated by λ, the Greek letter lambda. As the name suggests, *this is merely the length of a complete wave, measured from crest to crest* (A to B in Fig. 8-4) *or from trough to trough* (C to D).

There is a simple and very fundamental relationship between f, λ, and the speed of propagation of the waves v. Imagine yourself waiting in your car at a railroad crossing while a freight train passes by. Having nothing better to do, you find that 25 cars pass by per minute. And, being an old railroad man, you happen to know that each car is 42 ft long. So, if 25 cars, each 42 ft long, pass in a minute, $25 \times 42 = 1050$ ft of train pass per minute—and this is the train's speed: 1050 ft/min. So, transferring this analogy back to waves (or to any other moving periodic phenomenon), we discover that

$$v = f\lambda.$$

As an example, the note A above middle C on a piano should have a frequency of 440 vibr/sec. If the speed of sound (which depends to some extent on the temperature) is 1100 ft/sec, the wavelength of this sound is $\lambda = v/f = 1100/440 = 2.50$ ft.

8-5 Standing Waves

Suppose now that the far end of the rope in Fig. 8-4 is fastened to a wall, so that the train of waves is reflected backward from it. The rope will now be transmitting two wave trains simultaneously: the incoming hand-generated waves moving from left to right and the similar reflected wave train moving from right to left. To discover what will happen, let us first consider the situation in which a pair of single pulses traveling in opposite directions meet at some point along the rope.

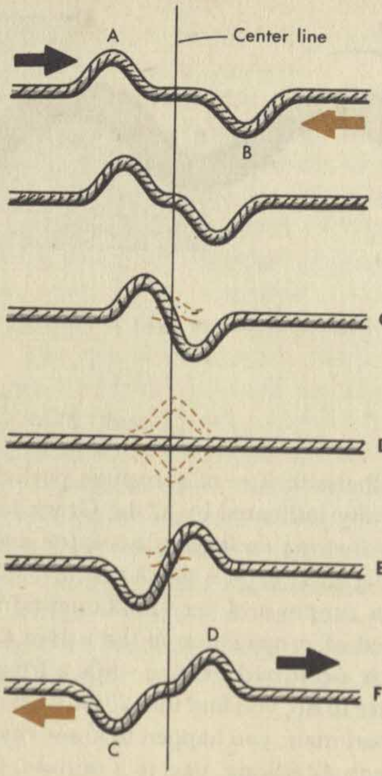


FIG. 8–5 Interference of two wave pulses passing each other on a rope.

An upward flick of the wrist will send a pulse along the top side of the rope, which be reflected back as a pulse on the lower side. As this reflection leaves the wall, another upward flick will send another pulse head-on toward the returning reflection. Figure 8-5A shows this situation, with *A* rushing to meet the reflection *B*.

A surprising property that waves of any sort generally have is their ability to pass through each other in the same medium without being changed or distorted in any way. This is shown in Fig. 8-5B to F, which sketches a series of instantaneous pictures of the collision, representing the shape of the rope at intervals only a small fraction of a second apart.

In Fig. 8-5A, the two pulses are approaching, and in Fig. 8-5B they are touching but have not yet interfered with one another. In Fig. 8-5C, the two pulses overlap, and the dashed lines show each pulse as it would have been *if it had proceeded undisturbed;* the solid line, indicating the rope itself, is the vector sum of the two interfering dashed-line waves. In Fig. 8-5D, the two pulses would have been exactly opposite one an-

other, and because they are of identical symmetrical shapes, the rope itself now lies in a perfectly straight line. We must not forget that, although the illustration shows the rope to be straight at this particular instant, the rope is nevertheless in motion. The particles to the left of the center line are moving downward, and those to the right of the center line are moving upward. The kinetic energy of the rope particles carries them an instant later to the configuration of Fig. 8-5E, which we can construct as before, the solid line of the rope being the sum of the dashed lines that show each pulse as it would have been if undisturbed by the other. In Fig. 8-5F, the two pulses, now separated again, go their independent ways just as though the collision had not occurred.

We should look back now over the six sketches and note that the center point of the rope on the center line, although two pulses have apparently been transmitted through it, has never stirred a hair's breadth from its original undisturbed position. How is it possible to transmit a wave along a rope if there is a piece of the rope in its path that does not move? We can get ourselves off the uncomfortably sharp horns of this dilemma by saying that the waves were *not* transmitted through the center point at all ! Pulse D in Fig. 8-5F is not pulse A (Fig. 8-5A) at all; instead, it is pulse B, which has been reflected as though the center line were a rigid wall. Similarly, C is not B passing through, but the reflection of A. This seems a reasonable explanation, since A's attempts to pull the center particle upward are exactly countered by B's simultaneous downward pulls; the result is that, for each wave, the center particle is as rigid and unyielding as though it were a solid wall.

The perfect reflections of Fig. 8-5 are possible only when the two waves are identical in shape and size. If A and B were not the same shape, the center point would move to some extent, so that C would be composed of a partial reflection of A, plus a portion of B which is able to sneak through. A very complicated mathematical analysis would show that C must nevertheless have exactly the same shape as B, so that although C's origin is something much more complex, it *appears* as though B had traveled through undisturbed.

A similar analysis will show the motion of a rope when it is carrying simultaneously two similar trains of waves moving in opposite directions. Figure 8-6A shows two sine waves: the light solid line represents the incoming wave moving from left to right, and the dashed line represents the reflected wave moving away from the wall from right to left. Note that neither of these lines represents the configuration of the rope; the displacement of the rope at every point is the *sum* of the displacements of the two waves. In Fig. 8-6A, these everywhere add up to zero, so the heavy line, representing the rope itself, is perfectly straight. (Some portions, however, have an upward velocity, and some a downward.) In Fig. 8-6B, the incoming wave has moved slightly to the right and the

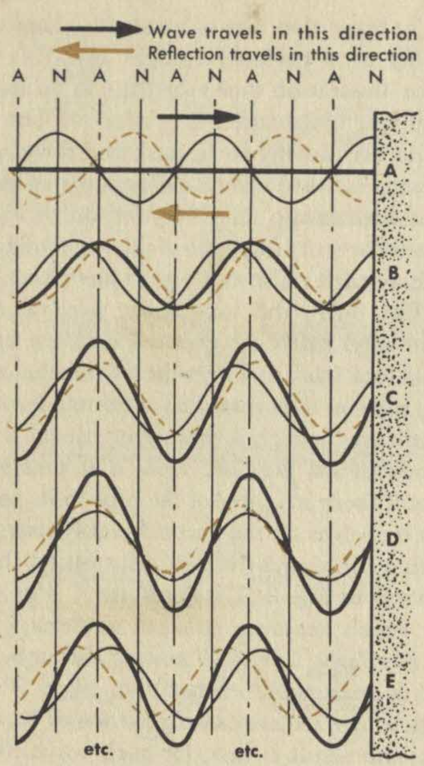


FIG. 8-6 Formation of standing waves in a rope or wire.

reflected wave slightly to the left. If we go along from point to point, again adding the displacements of the two waves, we get as their sum the wavy heavy line, which represents the rope a fraction of a second later than in Fig. 8-6A. The diagrams proceed at intervals of a fraction of a second, each one showing the incoming and reflected waves to have moved slightly from their previous positions, and the heavy line in each diagram representing the actual shape of the rope at that instant.

There are a number of points along the rope where the rope has no motion. These points are called *nodes*, marked N. From the drawing it is evident that the nodes are spaced a half-wavelength apart. Between the nodes are points where the amplitude of the motion of the rope is twice the amplitude of the generating waves. These points are called *antinodes* (A) and are, of course, also spaced $\lambda/2$ apart. Looking at a rope vibrating as shown in Fig. 8-6, we would see no motion along the rope, as we would with a single traveling wave; it would appear to be merely oscillating up and down between the nodes. For this reason, such a wave pattern is called a *standing wave*.

Since the rope, or wire, cannot move where it is fastened to the wall,

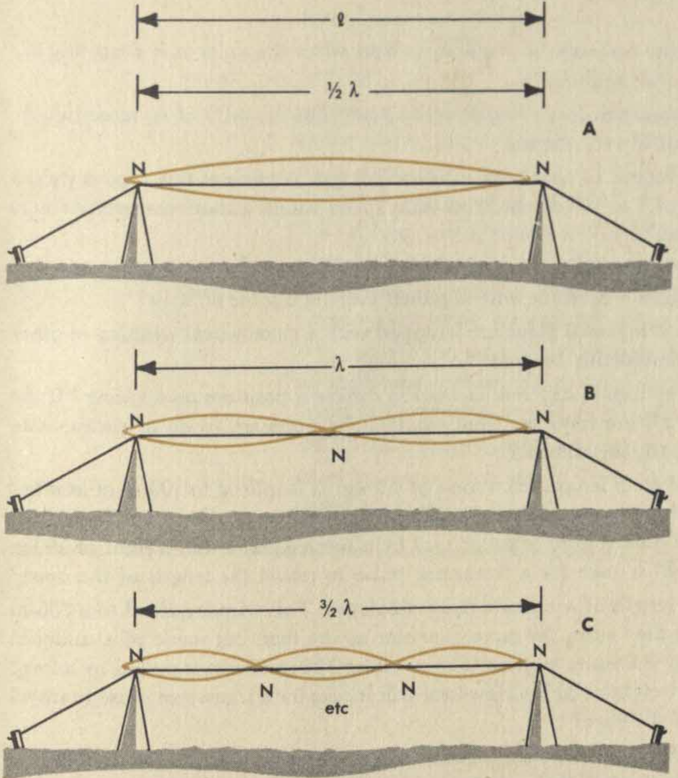


FIG. 8-7 A wire vibrating in its fundamental mode and in several overtones.

the wall must always be a node. And if, in Fig. 8-6, we clamp the rope rigidly at *any* node, the same standing wave will continue to vibrate between the clamp and the wall. This is the principle on which all stringed musical instruments operate.

Figure 8-7 shows a tightly stretched wire vibrating between two fixed points. When the wire is plucked, a hodgepodge of waves of all possible wavelengths are started along the wire. Most of these, since their reflections from the end points will not be in phase with one another, will fade out from interference with their own reflections. Their energy will not be lost, however, but will be transferred to those select vibrations whose wavelengths are just right to form nodes at the supports. The vibration of the longest wavelength (or lowest frequency, since $f = v/\lambda$ and v is the same for all the waves), shown in Fig. 8-7A, is called the *fundamental*; the other modes of vibration, such as those shown in Fig. 8-7B and C, are called the *overtones* of this fundamental.

In any stringed musical instrument, the string vibrates simultaneously in its fundamental mode, as well as in many of its overtones, or *harmonics*.

Questions

(8-1) 1. Analyze the equation $v = \sqrt{F/m_l}$ to find what the units of F must be, if v is in m/sec, and m_l in kg/m.

2. Analyze the equation $v = \sqrt{F/m_l}$ to show what the units of m_l must be, if v is in ft/sec and F is in pounds.

3. A 200-m length of cord has a mass of 500 gm. A piece of this cord is pulled with a force of 3×10^6 dynes. With what speed would a transverse pulse travel along the cord?

4. What is the velocity of a transverse wave along a wire, a thousand-foot coil of which weighs 8 lb, if the wire is pulled taut by a force of 20 lb?

5. The bass strings of a piano are wrapped with a close helical winding of other wire. Why should this be done?

6. Why are strings of different m_l used in different positions on a guitar? If the strings were all the same m_l, what would it be necessary to do to compensate when tuning the instrument?

7. A cord 50 m in length has a mass of 1.2 kg. It is spliced to 100 m of another cord made of the same material, but of only half as great a diameter. The resulting 150 m of spliced cord is pulled taut by a force equal to the weight of 40 kg. How long will it take for a transverse pulse to travel the length of the cord?

8. A 100-m length of wire has a mass of 800 gm. This wire is spliced to a 200-m length of another wire, the same diameter as the first, but made of a material whose density is 3 times as great. The resulting 300 m of wire is pulled by a force equal to the weight of 30 kg. How long will it take for a transverse pulse to travel the length of the wire?

(8-2) 9. In Question 7, if a transverse pulse is started at one end, two pulses should return. Explain why, and calculate when the two pulses should be expected.

10. In Question 8, if a transverse pulse is started at one end, two pulses should be expected to return. Explain, and calculate when the two pulses should be expected.

(8-4) 11. Light is a wave motion that travels at a speed of 3.00×10^{10} cm/sec. Its wavelength (which varies with color) can be accurately measured. What is the frequency of vibration of a green light that has a wavelength of 6×10^{-5} cm?

12. Radio waves are the same kind of electromagnetic waves as light and travel at the same speed of 3×10^{10} cm/sec. What is the wavelength of the radiation from radio station KBOL, which has a frequency of 1490 kilocycles/sec?

13. Sound waves with a frequency of 550 vibr/sec have (in air) a wavelength of 2 ft. What is the speed of sound waves in air?

14. A sound generator placed under water has a frequency of 725 vibr/sec, and creates waves of $\lambda = 2$ m. With what speed is a longitudinal pulse transmitted through water?

15. Cans of beans in a packing plant are placed 8 in. apart on a conveyor belt. In 1 min, 135 cans pass a stationary inspector. What is the speed of the belt, in ft/sec?

16. It has been said that a parade of Chinese twenty abreast could march forever past an observing point. If we assume the parade to march at 3 mi/hr, with ranks 4 ft apart, what is the Chinese birthrate, in babies/hr?

(8-5) **17.** A standing pattern of sound waves can be established in a glass tube, by reflection of the waves from pistons at each end. Cork dust in the tube will settle at the nodes, where the air is relatively undisturbed. If these nodes are measured to be 4 in. apart, (a) what is the wavelength of the sound? (b) What is its frequency?

18. A long wire is observed to be vibrating in the wind, and antinodes (where the wire is blurred by its motion) are seen to be 10 ft apart. At the same time, the wire emits a hum whose frequency is 50 cycles/sec. With what velocity are transverse waves being transmitted along the wire?

19. A tight guy wire 3.2 m long, with a linear mass of 4 gm/cm, is twanged at its midpoint and sounds a tone whose fundamental frequency is 20 vibr/sec. What is the tension on the guy wire?

20. A 12-ft chain weighing 0.96 lb/ft supports a weight of 5000 lb. When struck at its midpoint it vibrates in its fundamental mode. What is the frequency of the sound it emits?

chapter / nine

Sound

9-1 Sound Transmission

In the previous chapter, we looked into some of the properties of waves traveling along one-dimensional carriers such as ropes. Here, when we think of sound, we must think of waves spreading out in three dimensions from some vibrating source and transmitted as longitudinal waves through the air, or occasionally through water.

If we watch an artillery battery practice some distance away, we shall see the flash of light a few moments before we hear the sound of the shot. And we all know that the clap of thunder follows the lightning with a delay that depends on how far away the thunderstorm is. Since light propagates almost instantaneously (by our everyday standards, of course), the velocity of sound can easily be found by timing the lag between the flash of light and the roar of the shot and then measuring the distance between the observer and the artillery piece. In this way we find that ***the velocity of sound in air at 32°F is about 330 meters, or approximately 1100 feet, per second.***

As an example, we may imagine that we see a flash and flying debris

from an explosion on a mountainside some distance away. We hear the bang of the explosion 14.8 seconds later. A survey (or more easily, careful measurement on an accurate map) shows the location of the explosion to be 3.083 miles away. For the speed of the sound, then, we have

$$v = \frac{3.083 \times 5280}{14.8} = 1100 \text{ ft/sec.}$$

The velocity of sound in air is not appreciably affected by changes in pressure. We can qualitatively justify this experimental finding by comparing the transmission of the longitudinal waves in air with the way transverse waves are transmitted in a rope. If we increase the pressure on air (or any other gas), we may think of the resilience or "springiness" of the gas as being increased; this, we reason, should increase the speed with which a compression or rarefaction will be transmitted from one location in the gas to another location immediately adjacent, in the same way that increased tension speeds up the transmission of a transverse wave from one point on the rope to the next. However, there is a great difference between the effect of tension in the rope and pressure in the gas. An increased tension does not change the linear density of the rope enough to bother with. Air and other gases, however, are very easily compressible; as we shall see later, increasing the pressure on a gas reduces its volume and therefore increases its density in direct proportion to the increase in pressure. Increasing pressure thus tends to speed sound velocity by increasing the springiness of the gas; on the other hand, by making the gas more dense, it increases the inertia of the gas enough to completely nullify the gain. For this reason, a change in pressure has almost no effect on the speed of sound. An increase in temperature, though, tends to make a gas expand. Therefore, if the gas is heated, it must either expand and become less dense without changing the pressure, or, if the gas is confined so that it cannot expand, it will increase the pressure while the density remains the same. In either case, we see that sound must travel faster in warm gas than in cool.

A light gas, such as hydrogen, is much less dense than air is at the same pressure, and therefore sound travels much more rapidly in hydrogen than in air. Similarly, in a dense gas like carbon dioxide, the velocity of sound is slower than it is in air.

The factors of elasticity and density also govern the speed of sound in liquids and solids. A longitudinal compression wave along a metal rod, for example, is transmitted with a speed of

$$v = \sqrt{\frac{Y}{d}}$$

where Y is Young's modulus and d is the density of the material of which the rod is made.

For us to be able to speak of the *pitch*, or frequency, of a sound, the

sound must consist of a train of a considerable number of waves. One way to produce a train of waves of definite frequency is to force some solid surface to vibrate back and forth by mechanical or electrical means, thus causing alternate compressions and rarefactions which spread out through the surrounding air. The high note of a soprano that reaches your ear from the radio or TV is caused by the vibration of the loud-speaker diaphragm, which is forced to move back and forth thousands of times per second by the rapidly fluctuating pull of a magnet. A card or light stick held against the teeth of a rotating gear wheel will move back and forth as each tooth strikes it, and so set up in the air sound waves of the same frequency as its own motion.

The sound of nearly all musical instruments (and of our own voices as well) depends on setting up standing waves in strings (piano, guitar, violin, etc.) or in air columns (organ, flute, trumpet, etc.).

Figure 8-7 showed a wire vibrating in several different *modes*; vibrations of this sort are the basis for all stringed instruments. The fingering of a violin or guitar changes the length of the string included between the end nodes and thus changes the frequency of the fundamental vibration. Take, for example, a string 20 in. long whose tension is adjusted so that its fundamental vibration has a frequency of 440 vibr/sec. If a musician presses the string against a fingerboard, in effect reducing its length to 18 in., the fundamental wavelength has been reduced from 40 in. to 36 in. From our basic wave equation $v = f\lambda$, we see that the frequency must be increased in the same ratio as the decrease in λ. The shortened string will therefore vibrate with a fundamental frequency $f = 440 \times 40/36 = 489$ vibr/sec.

A vibrating string is so small that by itself it would create very weak waves in the surrounding air. For this reason, all stringed instruments are provided with some form of sounding board, a large flat surface on which the strings are supported and which vibrates in unison with the strings. Sounding boards respond more readily to some frequencies than to others; hence some overtones are emphasized and others are suppressed. It is this difference in the relative strengths of the various overtones that makes a guitar, for example, sound different from a banjo, and a zither different from a harpsichord.

Most wind instruments are in effect tubes that are closed at one end and open at the other; the effective length is controlled by valves that expose openings along the tube. Organ pipes are of two kinds—open at both ends, or open at one end and closed at the other. (A pipe or tube closed at both ends would not be very useful as a musical instrument—at least one end must be open, so that the waves of alternating compression and rarefaction can escape into the surrounding atmosphere.)

Figure 9-1 shows tubes of these two types, with standing air waves sketched inside them. The layer of air next to a closed end cannot move,

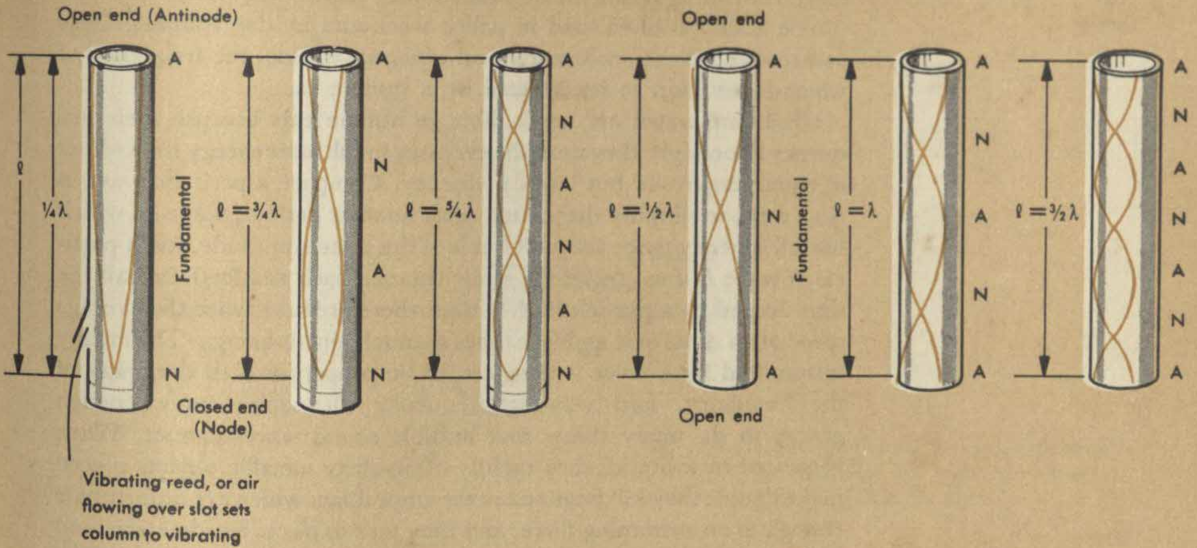


FIG. 9–1 The fundamental tone and some overtones in air columns vibrating in open-end and in closed-end tubes. (One end must always be open to allow the sound waves to escape.)

so a closed end must be a node; at the end open to the atmosphere, the vibrating air can move as freely as it is impelled to, so the open end must be an antinode. Like a plucked string or wire, the air column will vibrate at its fundamental frequency as well as in many overtones simultaneously. Figure 9-1 indicates only a few of the many possible overtones.

As an example, we can calculate the length of a closed-end organ pipe which is to produce a note whose fundamental frequency is 550 vibr/sec. (A closed-end pipe has one end closed, and the other open.) The sound will be traveling in air at a speed of about 1100 ft/sec. So, from the equation $\lambda = v/f$ we see that $\lambda = 2$ ft. Since the length of the tube is $\lambda/4$ (see Fig. 9-1), this gives us a required length of $\frac{1}{2}$ ft, or 6 in. (Actually this is not quite right; the antinode is not exactly at the open end but a slight distance beyond it, the distance depending on the wavelength and the tube diameter. For our purposes, we can neglect this refinement.)

9-2 Ultrasonics

The human ear cannot hear sound vibrations with a frequency of much less than 20 per sec or more than about 20,000 per sec. (As we get older, the high-frequency upper limit is reduced; many middle-aged people cannot detect frequencies higher than 10,000 per sec.) A dog's ear,

however, can hear sounds of considerably higher frequency, and this canine ability is often used in police work and by dog trainers. Dogs can receive "silent" orders from an *ultrasonic* whistle, the frequency of which is too high to be detected by a human ear.

Ultrasonic waves are not audible to human ears because their frequency is too high; they nevertheless carry much more energy than waves of equal amplitude but lower frequency. Compare a periodic wave A (in a rope, to simplify the matter) with another periodic wave B, which has a frequency twice that of A but is of the same amplitude. Each particle of wave B must travel the same distance back and forth in half the time needed by a particle in A. It must therefore have twice the average speed of an A particle and four times as much kinetic energy. The energy transmitted by a wave is thus seen to be proportional to the *square* of the frequency, and very-high-frequency ultrasonics carry enough energy to do many things that audible sound waves cannot. When generated in a liquid, they rapidly clean dirty metallic objects placed in the liquid; they kill frogs and other amphibians which are unfortunate enough to be swimming there; and they tear to pieces tiny bacteria and virus particles to provide biophysicists with an aid in their studies of the nature of life.

9-3 Supersonics and Shock Waves

Supersonics, which must not be confused with ultrasonics, is the study of the effects of objects that travel through a medium at a faster speed than the waves they generate. Nothing can move very fast through a solid, and even the most imaginative designers cannot yet dream of a submarine that travels faster than the speed of sound in water. So the practical problems of supersonics are essentially limited to planes and missiles that fly through the air at a speed greater than the speed of sound in air. In this situation, the moving object piles up disturbances—or waves—faster than they can get out of the way; this causes a single region of severe disturbance known as a *shock wave.*

The idea behind shock waves can be studied in a much commoner and much more gentle form by considering the bow wave of a ship traveling at a speed greater than the speed of the surface waves it generates. We see such a ship in Fig. 9-2. The ship, moving with speed v_s, has traveled from A to C during the time in which the surface wave it generated at A has moved from A to A' at a speed of v_w. Similarly, the wave started at B has at this same instant reached B', and so on for the infinite number of points that could have been drawn between A and C. The angle α between a bow wave and the direction of motion of the ship depends on the relative velocities of the waves and the ship:

$$\sin \alpha = \frac{A'A}{AC} = \frac{v_w t}{v_s t} = \frac{v_w}{v_s}.$$

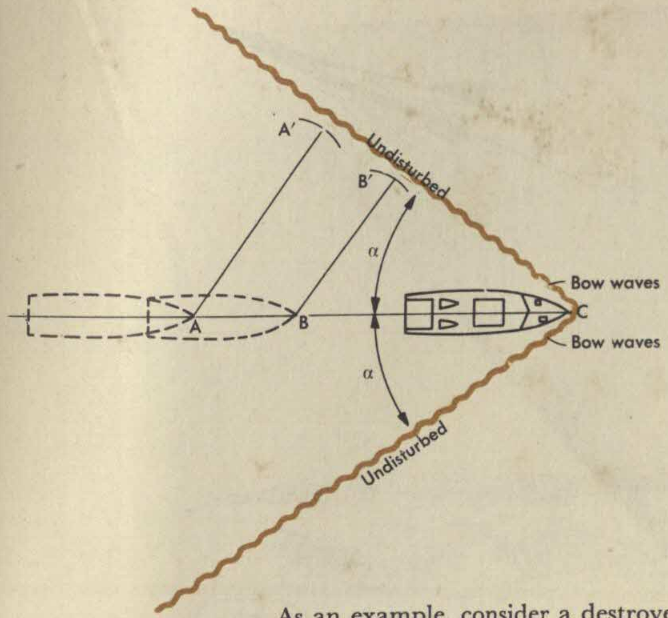


FIG. 9–2 The bow waves of a ship traveling faster than the speed of propagation of surface waves on the water.

As an example, consider a destroyer, cruising at 20 knots (1 knot = 1 nautical mile per hour = 1.151 miles per hour), and generating waves whose speed of propagation is 8 knots. We should expect the angle α between the path of the ship and the line of the "shock" bow wave to be given by

$$\sin \alpha = \frac{8}{20} = 0.400$$

$$\alpha = 23.6°.$$

A plane traveling faster than the speed of sound creates a similar disturbance in the air. Here, instead of the shock wave front forming a V on the water surface, it forms a giant cone in the air, since the waves generated by the passing plane spread out in all directions (Fig. 9-3). On the surface of the cone, where the waves pile up, there is a sharp pressure difference. When this cone (known as the *Mach cone,* after the physicist Ernst Mach) strikes a house, it sounds like a loud thunderclap, and may even be strong enough to shatter windows. These sonic booms are becoming familiar in many parts of the country where supersonic planes pass overhead.

The half-angle of the cone, α, is of course given by the same relationship that applied to the bow wave of a ship:

$$\sin \alpha = \frac{\text{speed of sound}}{\text{speed of plane}}.$$

Supersonic speeds are often given in terms of their *Mach number,* which is simply the ratio of the speed of the plane to the speed of sound. At Mach

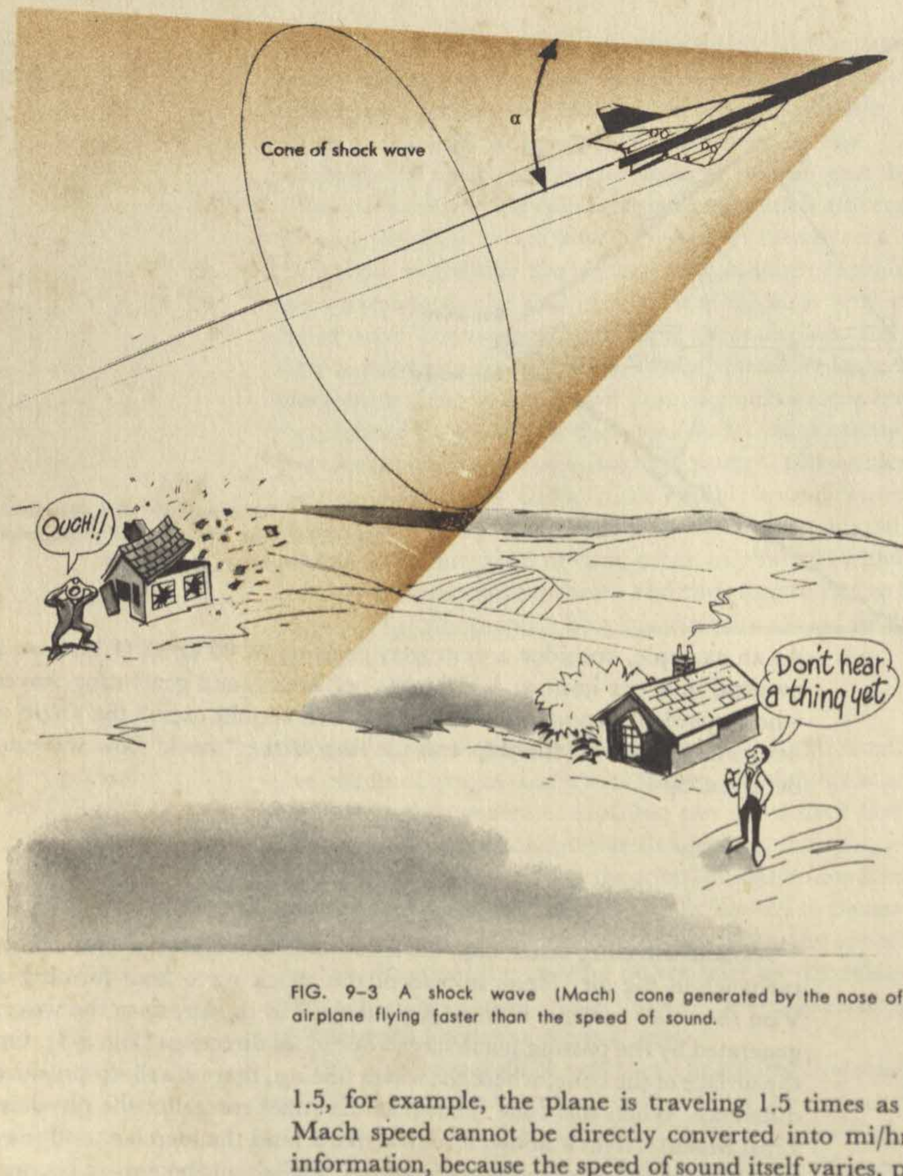


FIG. 9–3 A shock wave (Mach) cone generated by the nose of an airplane flying faster than the speed of sound.

1.5, for example, the plane is traveling 1.5 times as fast as sound. A Mach speed cannot be directly converted into mi/hr without further information, because the speed of sound itself varies, primarily with the temperature. On a pleasant summer day near the earth's surface, Mach 1 (the speed of sound) may be about 750 mi/hr; 30,000 ft overhead, the temperature may be $-70°F$, and at this cold temperature Mach 1 would be about 640 mi/hr. If the speed of a plane or missile is given in Mach units, it follows that $\sin \alpha = 1/\text{Mach number}$.

9-4 The Doppler Effect

Everyone has experienced (although many may not have observed) the Doppler effect. If a car, which is sounding its horn, rapidly approaches you, passes, and then continues on, a very noticeable drop

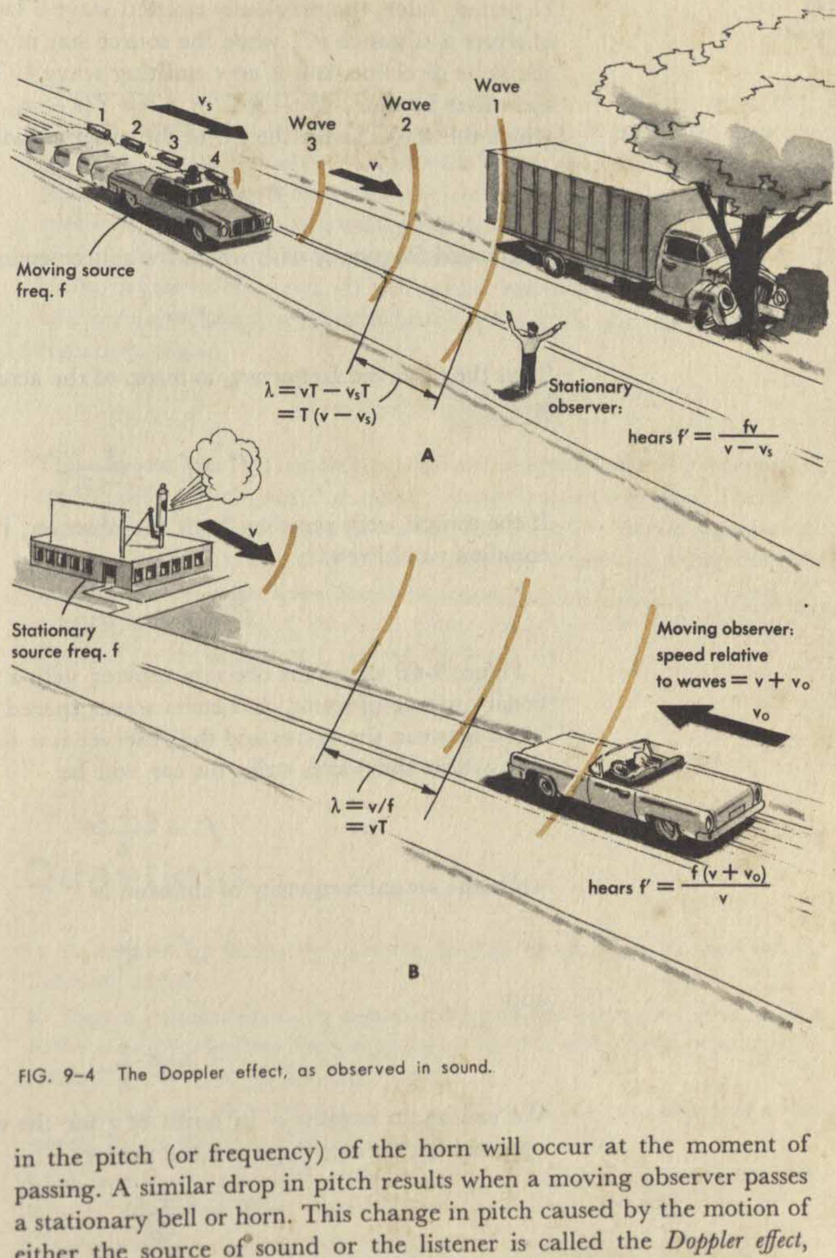


FIG. 9-4 The Doppler effect, as observed in sound.

in the pitch (or frequency) of the horn will occur at the moment of passing. A similar drop in pitch results when a moving observer passes a stationary bell or horn. This change in pitch caused by the motion of either the source of sound or the listener is called the *Doppler effect*, after the nineteenth-century Austrian physicist Christian Doppler.

Let us look first at a moving source. The top of Fig. 9-4A, at $t = 0$, shows a source of sound S in the process of emitting a wave. The source is moving toward the observer with a speed v_s. At a time T seconds

(1 period) later, the previously emitted wave 3 has traveled toward the observer a distance vT, while the source has moved a distance v_sT in the same direction and is now emitting wave 4. The distance between the waves is, then, $vT - v_sT$, or $\lambda' = T(v - v_s)$. The frequency with which the waves strike the ear of the observer will thus be

$$f' = \frac{v}{\lambda'} = \frac{v}{T(v - v_s)}.$$

The actual frequency with which the source emits waves is

$$f = \frac{1}{T}.$$

Thus the observed frequency, in terms of the actual emitted frequency, is

$$f' = f \times \frac{v}{v - v_s}.$$

If the source were receding from the observer, it is apparent that the equation would read

$$f' = f \times \frac{v}{v + v_s}.$$

Figure 9-4B shows the observer moving with a speed v_o toward a stationary source of sound that emits waves spaced λ apart. The relative speed between the waves and the observer is $v + v_o$, and the frequency with which the waves strike his ear will be

$$f' = \frac{v + v_o}{\lambda}$$

while the actual frequency of emission is

$$f = \frac{v}{\lambda}$$

and

$$\lambda = \frac{v}{f}.$$

We can again express f' in terms of f for the observer who is either approaching or receding from the source:

$$\begin{aligned} f' &= \frac{v \pm v_o}{v/f} \\ &= f \times \frac{v \pm v_o}{v}. \end{aligned}$$

The above equations can be combined into the following single expression:

$$f' = f \times \frac{v \pm v_o}{v \pm v_s}.$$

There are rules governing the selection of the + and − signs, but a smattering of common sense will make these rules unnecessary. Approach (of either source or observer) makes the heard frequency higher, and the proper sign can be chosen to make the equation fit the situation. For example, consider car A speeding down the highway at 90 ft/sec. Ahead of A is car B, going in the same direction at only 30 ft/sec, and A sounds his horn before he passes. If the actual frequency of A's horn is 1000 vibr/sec, what frequency does B hear? Let us write our equation without regard for sign:

$$f' = 1000 \times \frac{1100 \pm 30}{1100 \pm 90}.$$

The source (car A) is moving toward the observer (car B); this motion, considered by itself, would make the heard frequency greater. Therefore, we must make the denominator of the fraction (which contains v_s) smaller by subtracting the 90. The observer is headed away from the source; this motion, considered by itself, would make the heard frequency less. Therefore, we must make the numerator (which contains v_o) smaller by subtracting the 30. So we have

$$f' = 1000 \times \frac{1100 - 30}{1100 - 90}$$
$$= 1059 \text{ vibr/sec.}$$

Questions

1. A thunderclap follows the lightning flash by 4.0 sec. How far away did the lightning strike?

2. Derive a simple formula by means of which you can count the time in seconds between a lightning flash and the sound of thunder, and from it determine the distance in miles to the lightning stroke.

3. A man stands 400 ft in front of a high cliff and fires a gun. How long a time elapses before he hears the echo reflected from the cliff?

4. A man stands some distance from a high cliff and fires a gun. He hears the echo 1.5 sec after the gun was fired. How far away is the cliff?

5. If you stand at one end of a long hallway with a good sound-reflecting surface at its far end, you can clearly hear the echo of a hand clap. With a little effort and practice, you can clap your hands rapidly and regularly, so that each clap coincides with the echo of the previous one. Assume you stand 135 ft from the end of the hall, and a colleague counts 120 claps in 30 sec. What is the speed of sound in this hall?

6. Through binoculars a man watches a carpenter driving nails at a regular rate of 1 stroke/sec. He hears the sound of the blows exactly synchronized with the blows he sees. He hears two more blows after he sees the carpenter stop hammering. How far away was the carpenter?

7. A steel rod (density 7.8 gm/cm^3) 50 m long is struck on the end with a hammer. How long before the reflection of the longitudinal pulse returns?

8. An aluminum pipe (density 2.7 gm/cm^3) 75 m long is struck on the end with a hammer. How long before the reflection of the longitudinal pulse returns?

9. A siren can be made from a rotating flat disk, which has regularly spaced holes punched through it along a circle concentric with the axis of rotation. An air nozzle is directed against the disk, and each time a hole passes the nozzle, a puff of air is released to generate a wave pulse. What frequency sound will be produced by a disk containing 72 holes, and rotating 1800 rev/min?

10. A strip of stiff plastic is held against the teeth of a rotating 48-tooth gear, and the sound produced has a frequency of 512 vibr/sec. What is the rotational speed of the gear, in rev/min?

11. A man stands some distance in front of a long, high flight of stairs, and fires a gun. The stairs have treads (the horizontal surfaces of the steps—the surface one steps on) that are uniformly 15 in. deep. Will there be any tone associated with the echo from the steps? If so, what is its frequency?

12. A man in front of a long flight of stairs makes a loud clap by slapping two boards together. The tread of the stairs is 11 in. What frequency is associated with the sound reflected from the stairs?

13. A string on a musical instrument is 33 in. long, and has a fundamental frequency of 192 vibr/sec. What is the frequency when this string is pressed against a fret to reduce its length to 30.5 in.?

14. A string on a musical instrument is 22 in. long, and has a fundamental frequency of 272 vibr/sec. To what length must the string be shortened (by pressing it against a fret or a fingerboard) to make its frequency 318 vibr/sec?

15. (a) What is the fundamental frequency of a tube open at both ends and 12 in. long? (b) What is the frequency of its first overtone? (c) of its second overtone?

16. (a) What is the fundamental frequency of a tube closed at one end and 8 in. long? (b) What is the frequency of its first overtone? (c) of its second overtone?

17. The fundamental frequency of a piano string tuned to the note A above middle C is 440 vibr/sec. What is the frequency of its first overtone? Of its third overtone?

18. The fundamental frequency of an organ pipe open at both ends is 440 vibr/sec. What is the frequency of its first overtone? Of its third overtone?

19. What is the length of the organ pipe of Question 18?

20. What is the length of an organ pipe closed at one end, having a fundamental frequency of 440 vibr/sec?

(9-2) 21. Take two "sounds" of the same amplitude, one of 2000 vibr/sec and the other an ultrasonic one of 100,000 vibr/sec. How many times as much energy does the latter carry than the first?

22. A mechanism for creating 3000 cycle/sec sound waves of a certain amplitude in water requires a power of 25 watts. How many watts would be needed to create waves of the same amplitude, but with a frequency of 60,000 cycles/sec?

(9-3) **23.** A ship moving at 30 knots creates bow waves including an angle of 60° between them. What is the speed of the surface waves created by the ship?

24. A toy sailboat on a pond moves with a speed of 1 m/sec and creates waves whose speed is 30 cm/sec. What is the angle included between its bow waves?

25. A bullet travels 510 m/sec through the air. What will be the angle between the shock wave and the path of the bullet?

26. A jet plane travels at Mach 2 (2 times the speed of sound) at an altitude of 5000 ft. How far past an observer will the plane be when the shock wave hits him?

(9-4) **27.** At what speed would a car have to be approaching an observer for the latter to hear music from the car radio with a pitch 10 percent higher than it actually is?

28. At what speed would an observer in a car have to approach a stationary siren in order that its pitch will sound 10 percent higher than it actually is?

29. A car has a siren sounding a 2000-vibr/sec tone. What frequency will be heard (a) by a stationary observer as the car approaches him at 45 mi/hr? (b) If the car is stationary and the observer approaches it at 45 mi/hr? (c) If the car is going away at 60 mi/hr and the observer chases it at 30 mi/hr?

30. A car, sounding a horn whose frequency is 1200 vibr/sec, drives directly toward a vertical cliff at 55 ft/sec. What is the frequency of the sound of the horn reflected from the cliff, as heard by the driver of the car? (The car acts as a moving source when the sound is emitted, and as a moving observer for its reflection.)

chapter / ten

Temperature and Heat

10-1 Measurement of Temperature

Except for judging by the feel of our skin, which is not a very accurate quantitative guide and is often deceiving, the only way we can measure temperature is to measure the effects temperature changes have on the physical properties of materials. All physical bodies, such as a piece of solid material or a certain amount of liquid, nearly always respond to temperature changes in a simple way: they expand when the temperature rises and contract when it drops. There are many other effects, and most of them are utilized in one way or another to measure temperature: the change in the electrical resistance of wire; the variation with temperature of electric current generated by dissimilar metals joined together; changes in the viscosity or plasticity of materials; changes in the color of the light emitted by very hot bodies, and so on.

The most commonly employed effect, however, is that of expansion. Thermometers used for this purpose are usually constructed in such a way that a small thermal expansion results in a large displacement of an indicator. In thermometers based on the expansion of a liquid, a

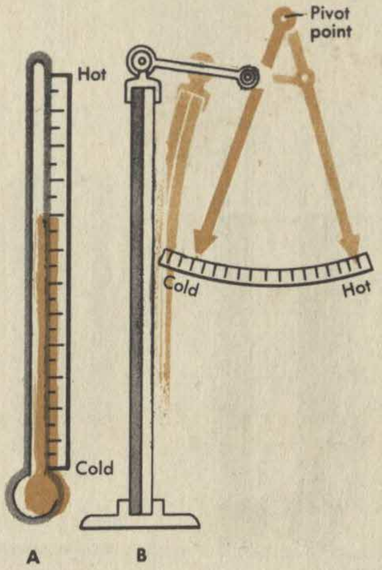


FIG. 10–1 (A) Mercury thermometer; (B) bimetal strip thermometer.

comparatively large amount of the liquid (usually mercury or alcohol) is confined in a bulb, and its expansion causes the excess liquid to rise in a narrow capillary tube (Fig. 10-1A). In thermometers based on the expansion of solids, a double (bimetallic) strip is often used, composed of two metals whose expansions are different when they are heated. For example, in Fig. 10-1B, the darkened strip might be of brass, and the unmarked strip (cemented or otherwise securely fastened to the brass strip) could be zinc. Brass expands considerably more than zinc when heated, so as the bimetallic strip is warmed, the greater expansion of the brass forces the strip to curve, as shown, and move an indicator across a scale.

Whatever kind of thermometer we choose, the scale on which the indicator moves must be calibrated in some definite, easily reproducible manner. The two thermometric scales in general use are calibrated at the freezing point and the boiling point of water (at a pressure of 1 standard atmosphere). On the *Centigrade*, or *Celsius*, scale, which is used in scientific work all over the world, the freezing point is 0°C, and the boiling point is 100°C. On the *Fahrenheit* scale, used (except for scientific work) in the United States, 32°F is the freezing point, and boiling point is 212°F. Figure 10-2 shows the calibration of a pair of thermometers. With both thermometers immersed in a beaker of wet crushed ice (Fig. 10-2A), we mark the points where the mercury stands as 32 on the Fahrenheit thermometer, and as 0 on the Centigrade. Similarly,

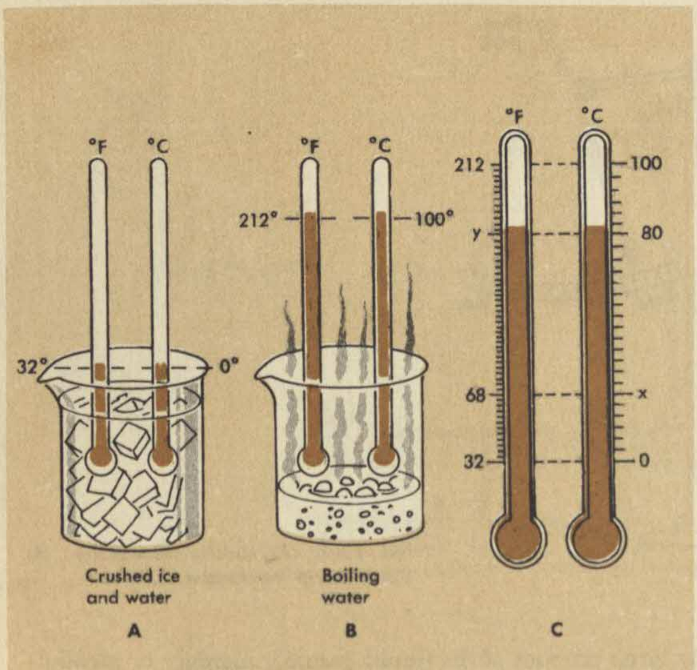


FIG. 10–2 Calibration of Fahrenheit and Centigrade (or Celsius) thermometers.

with the thermometers in steam above boiling water (Fig. 10-2B), we mark the points 212 and 100. On the Fahrenheit scale the interval between freezing and boiling is thus divided into 180 degrees (212–32), and on the Centigrade scale this same interval is divided into 100 degrees. From this it is plain that the Fahrenheit degree is only $\frac{5}{9}$ the size of the Centigrade degree.

Any definite Fahrenheit temperature corresponds to some definite Centigrade temperature, of course, and vice versa. There are formulas for these conversions, which both authors of this book have gone to great pains *not* to learn. They are too easily forgotten, or remembered wrong; it is considerably more dependable to reason from scratch on each conversion. For example, in Fig. 10-2C, the temperature 68°F is shown. This temperature is $68 - 32 = 36$ Fahrenheit degrees above freezing, which is $36 \times \frac{5}{9} = 20$ Centigrade degrees above freezing. Since the freezing point is 0°C, this means that 68°F is the same as 20°C. For a conversion in the other direction, the temperature 80°C is shown. Here we find 80 Centigrade degrees above freezing, or $80 \times \frac{9}{5} = 144$ Fahrenheit degrees above freezing. Since freezing is 32°F, 80°C must be $144 + 32 = 176$°F.

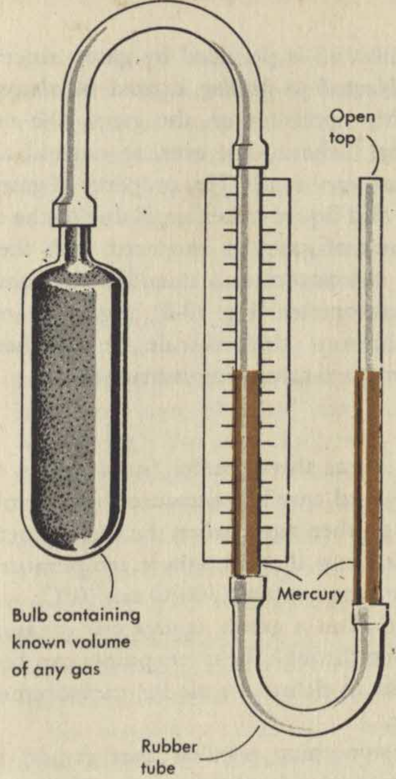


FIG. 10–3 A gas thermometer. The changing volume of the trapped gas is read from the height of the mercury column in front of the scale. In order to keep the mercury level the same on both sides, the right-hand tube is lowered as the gas heats; the gas in the bulb will therefore remain at atmospheric pressure at all temperatures.

10-2 Gas Thermometers

Although all thermometers will behave, generally speaking, in a similar way and give us a clear indication of whether the temperature goes up or down, they will disagree among themselves in smaller details, since different materials react somewhat differently to an increase of temperature. Thus, if we mark by 0 and 100 the freezing and boiling points of water on thermometers filled with mercury, alcohol, and water, we shall find that, as the temperature increases from zero, the mercury and alcohol columns will rise while the water column will first drop and will begin to rise only after the other two columns have covered about 4 percent of the total distance to the boiling point. Even the mercury and alcohol columns, which are adjusted to show identical values at the two ends of the scale, will not quite agree with each other in between, because they expand at different rates in different temperature intervals.

The physiological sense of heat is too vague to be used for any exact definition of temperature, and we have no reason to give preference to any particular one of the liquid thermometers filled with mercury, alcohol, water, or olive oil; or to a particular solid one formed by iron-copper, silver-zinc, or other bimetallic strip. We must therefore look somewhere else for an exact and universal definition of the temperature

scale. The solution is provided by gases, since it has been found that ***all gases subjected to heating expand in almost exactly the same way.*** The lower the pressure on the gases, the more nearly alike do all kinds of gases behave, but even at normal atmospheric pressure the differences are very small. This property of gases, which is in contrast to that of solid and liquid materials, is due to the extreme simplicity of the inner structure of gases as compared with the structure of solids and liquids. We can accept as a standard the temperature scale provided by a gas thermometer (Fig. 10-3), regardless of what gas is used to fill it. Having this as a standard scale, we can then properly calibrate any other temperature-measuring instrument.

10-3 Absolute Zero

Let us take a gas thermometer (such as the one shown schematically in Fig. 10-3) and carefully measure the gas volume, first at the boiling point of water, then again when the thermometer is in a slush of ice and water. In defining the Centigrade temperature scale, we have agreed to call these temperatures 100°C and 0°C; we can now plot them as points B and F on a graph against our measured volumes (Fig. 10-4). A straight line through these two points can be drawn, and extended to left and right to define a scale for measurement over a wide range of temperature.

There is something peculiar that should be noted about the left (low-temperature) side of the graph. We cannot extend the line inde-

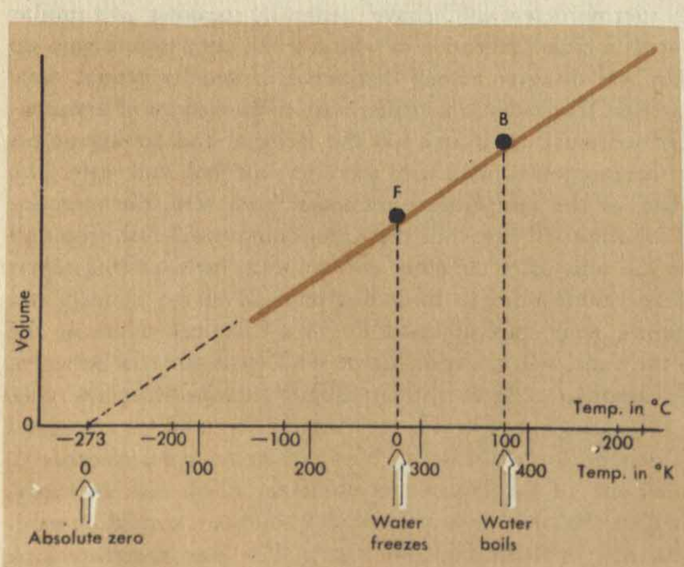


FIG. 10–4 The behavior of a gas at low temperature.

finitely, as we can on the other (high-temperature) side, because it soon runs into the axis of the graph, indicating a zero volume. It would be ridiculous to extend the line any farther, since as far as we know, a gas with a *negative* volume is an idea that has no meaning. So this point, toward which our graph seems to be heading, is called the *absolute zero* of temperature.

The apparent intention of the gas to shrink to zero volume at absolute zero is naturally never fulfilled. All gases liquefy before this point is reached; in fact, even before it begins to liquefy, a gas commences to deviate considerably from its more regular behavior at higher temperatures. Ingenious experimental procedures and theoretical corrections, however, have revealed that the temperature of absolute zero is −273.15°C. We shall not worry ourselves about the odd hundredths of a degree and shall be content to refer to this ultimate coldness as −273°C. A temperature scale beginning with 0 at absolute zero is an *absolute temperature scale*. The most common absolute scale uses the same size degree as the centigrade scale and is called the *Kelvin* scale (°K), named for Lord Kelvin, a great nineteenth-century English physicist. Since the Kelvin degree is the same size as the Centigrade degree and begins counting 273° lower on the scale, it is apparent that in order to convert a temperature given in °C into °K, we need only add 273.

The concept of an absolute zero is an important one in physics, and we shall later try to interpret it in terms of something more significant than the crossing point of two lines on a graph.

10-4 Pressure in Gases

As was apparent in the discussion of gas thermometers, there are three factors of importance in determining the behavior of gases: temperature, volume, and pressure. We have defined several temperature scales, and the meaning of volume is self-evident. At this point, however, we can well afford to devote a few paragraphs to a closer look at pressure and its effect on gases in general.

Although we seldom pay much attention to it, the air around us and in our lungs and ears is under a very considerable pressure. Earth's atmosphere is a mixture of about 78 percent nitrogen, 21 percent oxygen, 1 percent argon, 0.03 percent carbon dioxide, plus smaller amounts of many other gases, and a variable amount of water vapor. The total mass of the atmosphere is about 5×10^{15} tons, which amounts to more than one kilogram for every square centimeter of the earth's surface. The weight of the air above our heads was first measured by an Italian physicist, E. Torricelli (1608–1647), using an arrangement shown schematically in Fig. 10-5. He took a long vertical glass tube that was closed on the top, completely filled it with water, and placed it in a water-filled tub. When he opened a faucet at the bottom of the tube, the water level in

FIG. 10–5 Torricelli's experiment.

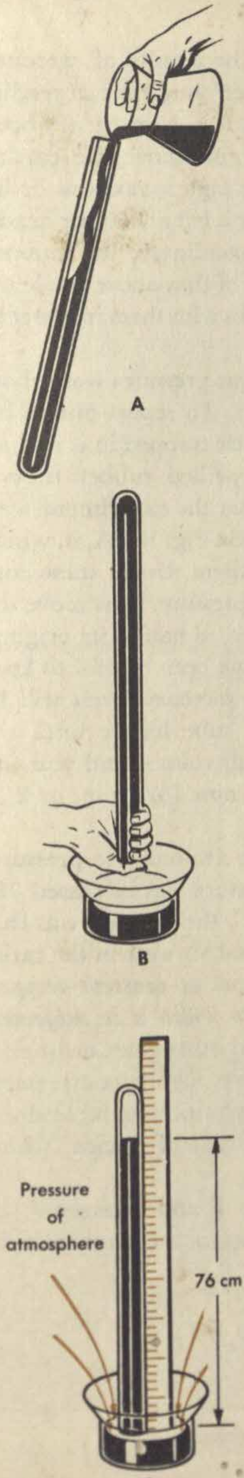


FIG. 10–6 The mercury barometer.

the tube fell to a height of about 10 m above the water level in the tub, leaving "Torricelli's emptiness"—or *vacuum*, as we call it now—at the upper part of the tube.

Since the volume A in Fig. 10-5 contains nothing but a little water vapor and is completely separated from the atmosphere, it can exert no appreciable force downward against the top of the column of water. So the atmospheric pressure at B, the surface of the water in the tub, is able to force water up the tube without any resistance except that resulting from the height of the water column itself. Hence the column of water (or any other liquid) can be used to measure the pressure of the atmosphere:

$$h \times d = \text{atmospheric pressure}$$

where h is the height of the fluid column and d is its density.

In our version of Torricelli's experiment, with the water standing exactly 10 m above the water surface in the tub, the pressure of the atmosphere in Italy that day must have been

$$\begin{aligned} 1000 \text{ cm} \times 1 \text{ gm/cm}^3 &= 1000 \text{ gm/cm}^2 \\ &= 9.80 \times 10^5 \text{ dynes/cm}^2. \end{aligned}$$

This experiment of Torricelli can be repeated in the laboratory much more conveniently by using mercury instead of water.* Figure 10-6 shows how to do this. A glass tube, closed at one end, is filled *completely* with mercury (Fig. 10-6A); put your thumb over the open end, invert the tube, and submerge the open end (thumb and all) in a dish of mercury (Fig. 10-6B). If you are at sea level and the atmospheric pressure is neither abnormally high nor low at the time, when the thumb is removed, the mercury column will fall until its top is about 76 cm above the mercury surface in the dish (Fig. 10-6C). At this height, the pressure at the bottom of the column is exactly balanced by the pressure of the atmosphere against the open surface of the mercury. This simple device is called a *barometer*, and the pressure of the atmosphere is often referred to as *barometric pressure*. Although it is not an actual pressure unit (which must have the dimensions of force/area), the height of a mercury column is such a convenient way of measuring pressure that pressures are often given in terms of "centimeters of mercury" or "millimeters of of mercury" (cm Hg or mm Hg, Hg being the chemical symbol for mercury).

Scientists all over the world agree, for example, to use exactly 76 cm Hg as the pressure of a *standard atmosphere*, which is very nearly the aver-

* Since mercury is 13.6 times denser than water, the mercury column will stand only 1/13.6 as high as Torricelli's column of water.

age atmospheric pressure at sea level. Using the density of mercury ($13.6\ gm/cm^3$) and the average value of $980\ cm/sec^2$ for g, we can readily find that one standard atmosphere is equivalent to a pressure of about $1.013 \times 10^6\ dynes/cm^2$, or $14.7\ lb/in.^2$. As we rise above the earth's surface (or rather, above sea level) by climbing high mountains or by ascending in a balloon or airplane, less and less air is left above our heads, and the atmospheric pressure decreases correspondingly; the amount of air over the top of Mt. Everest is only a third of that above sea level, and climbers on it are forced to carry oxygen tanks with them in order to breathe.

Among the earliest accurate experiments on gas pressures were those of the Irish physicist Robert Boyle (1627–1691). To repeat one of his experiments, take a closed glass tube with some air trapped in it and an open tube, and connect them with a mercury-filled rubber tube a couple of meters long, as shown in Fig. 10-7. Start the experiment with the two glass tubes in the relative positions shown in Fig. 10-7A, in which the mercury stands at the same level in both of them. Under these conditions the trapped air is at normal atmospheric pressure. Now move the open tube up until the trapped air is compressed to half of its original volume. You will find that in this case, if you have been careful to keep the temperature constant, the difference in the mercury levels will be about 760 mm (Fig. 10-7B).* Move the open tube higher until the trapped air is squeezed into a third of its original volume and you will find that the difference in the mercury levels is now 1520 mm, or 2×760 mm (Fig. 10-7C).

In Fig. 10-7A, the trapped air was subject to atmospheric pressure, i.e., 760 mm of mercury. In Fig. 10-7B, the pressure was increased 760 mm to a total of two atmospheres. In Fig. 10-7C, the pressure was that of three atmospheres. Since the volumes of trapped air were in the ratios $1 : \frac{1}{2} : \frac{1}{3}$, we can conclude that ***the volume of gas at constant temperature is inversely proportional to the pressure to which it is subjected,*** which is the classical *Boyle's law of gases.* Air and other gases ordinarily follow this law quite well, but when they are highly compressed, deviations from it can be observed. This is quite understandable since, in these cases, the density of gas approaches that of liquids, which possess very low compressibility.

The inverse proportionality between volume V and pressure P for a given sample of gas whose temperature does not change can be stated mathematically as

$$\frac{P_1}{P_2} = \frac{V_2}{V_1} \quad \text{or} \quad P_1V_1 = P_2V_2.$$

* It will vary slightly, depending on the atmospheric pressure.

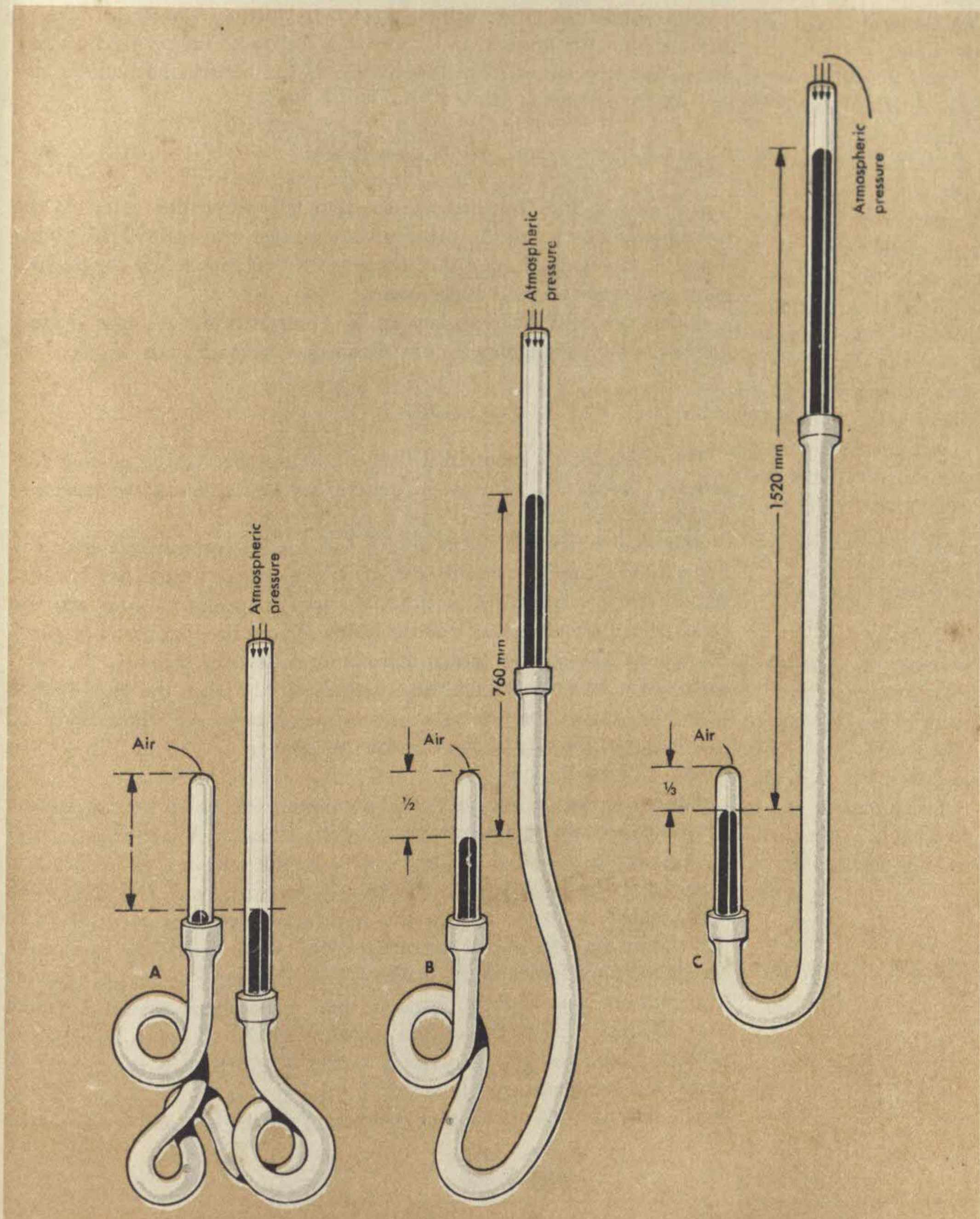


FIG. 10–7 An apparatus for determining the compressibility of gases.

10-5 The General Gas Law

The relationship between the volume and the temperature of a gas under constant pressure was investigated by Jacques Charles of France in the eighteenth century. Charles's work indicated the proportionality shown by the graph in Fig. 10-4, which we can express mathematically as

$$\frac{V_1}{V_2} = \frac{T_1}{T_2}.$$

Referring to Fig. 10-4, we can see that this proportion refers to the corresponding sides of the two similar triangles drawn from the points F and B, for example, and that ***the proportion is true only if the temperature used is the absolute temperature.***

Boyle's law and Charles's law can be combined into a single general gas law, which states that for any given and constant mass of gas,

$$\frac{P_1V_1}{T_1} = \frac{P_2V_2}{T_2}.$$

We have already mentioned that the *temperatures must be absolute temperatures*; before we try to use this general gas law equation, we must also emphasize that the *pressures must be absolute pressures.*

Everyone is accustomed to seeing and reading pressure gauges; tire gauges and gauges on compressed-air tanks or water tanks are familiar sights. When you have a blowout, the gauge applied to your flat tire reads zero. Does this mean that the inside of your tire is at zero pressure? No, not at all—it only means there is no difference between the pressures inside and outside the tire. In other words, when the gauge reads zero, the air in the tire is at atmospheric pressure, which may be anything from nearly 15 lb/in.^{2} near the seashore to 10 lb/in.^{2} in a town in the high mountains.

Let us use the gas law on a problem about tires, which we can assume keep the same volume no matter what the pressure. A man fills his tires to a gauge pressure of 28 lb/in.^{2} early in the morning ($T = 5°C$) in a city at sea level (atm. pr. = 15 lb/in.^{2}). By afternoon ($T = 25°C$), he has reached a town high in the mountains (atm. pr. = 10 lb/in.^{2}). He checks his tires in this mountain town. What should he expect the gauge to read? Down in the sea-level city, the gauge indicated that the tire pressure was 28 lb/in.^{2} greater than atmospheric pressure, which was 15 lb/in.^{2}. Thus the *absolute* pressure in the tires was $28 + 15 = 43$ lb/in.^{2}, the temperature was $5 + 273 = 278°K$. Up in the mountains, the temperature of the air in the tires is $25 + 273 = 298°K$. The unknown (but unchanging) volume of the tire is V, and we can write

$$\frac{43V}{278} = \frac{P_2V}{298}$$

or

$$P_2 = 43 \times \frac{298}{278} = 46.1 \text{ lb/in.}^2 \text{ (absolute)}.$$

What the gauge will indicate is the difference between this 46.1 lb/in.^{2} inside the tires and the 10 lb/in.^{2} of the outside atmosphere, or 36.1 lb/in.^{2}.

Actually, no real gas exactly follows the general gas law. Indeed, one way to define the physicists' imaginary "ideal gas" is to say it is one that *does* follow this law exactly. Real gases deviate from it for two reasons. If we look at the law, we see that if the pressure is raised higher and higher (and the temperature is kept constant), the volume must get smaller and approach zero for enormously high pressures. A real gas cannot do this, because the gas molecules themselves occupy space. Another factor which enters into the behavior of a real gas is the fact that the molecules actually attract one another slightly, especially when the gas is compressed and the molecules are close together. This attraction helps to squeeze the gas and makes its volume at high pressures a tiny bit smaller than it would be if it were an ideal gas. If one is using a gas thermometer to do very accurate thermometry, these small deviations must be taken into account. However, we do not need to aspire to this hundredth-of-a-degree accuracy in our work here and so may safely assume that the gases we deal with follow the general gas law.

10-6 Expansion of Solids and Liquids

In contrast to the gases, all of which expand by almost the same amount when heated, solids and liquids are individualists. The amount a solid material will expand when heated is measured by its *coefficient of linear expansion*, generally designated by α, the Greek letter *alpha*. This coefficient gives the fraction by which the length (or width or thickness) of an object will increase for each centigrade-degree rise in temperature. The total increase in length (Δl) that a body experiences will accordingly be the coefficient α multiplied by the length l, times the number of degrees rise in temperature (ΔT). In an equation,

$$\Delta l = \alpha l \Delta T.$$

Here are the linear coefficients of expansion for a number of materials:

Aluminum	$\alpha = 25 \times 10^{-6}/°\text{C}$
Brass	18×10^{-6}
Ice	50×10^{-6}
Invar*	0.9×10^{-6}
Steel	11×10^{-6}
Platinum	9×10^{-6}
Glass	9×10^{-6}

* Invar is a special nickel-steel alloy designed for use where a low coefficient of expansion is needed.

As an example, consider a steel bridge 200 m long, in a locality where the temperature may vary from $-30°C$ in the winter to a blistering summer heat of $+40°C$. This is a temperature change of 70°C, and from winter to summer the bridge will lengthen by an amount given by

$$\Delta l = 11 \times 10^{-6} \times 2 \times 10^{4} \times 70 = 15.4 \text{ cm}.$$

To allow room for such expansions, many bridges have one end mounted on rollers or some other device which will permit it to expand and contract freely with the changing seasons.

The table above shows that glass and platinum have equal coefficients of expansion. For this reason, platinum wire is often used, in spite of its expense, when the wire must pass through the wall of a piece of glass equipment. Since the two materials will expand and contract by the same amount, there is no danger of the glass cracking, or the wire shrinking from it as the temperature changes.

If the length and width and thickness of a piece of material increase as it is heated, its volume will also increase. The Greek letter *beta* (β) is often used to represent the *volume coefficient of expansion*, which is the fraction of its volume by which a piece of material will expand per centigrade-degree rise in temperature. Consider a rectangular block of material $a \times b \times c$ in size, which has a linear coefficient of expansion α. If we raise the temperature of the block by 1°C, its new dimensions will be $(a + \alpha a) \times (b + \alpha b) \times (c + \alpha c)$, or $a(1 + \alpha) \times b(1 + \alpha) \times c(1 + \alpha)$. The volume becomes $abc(1 + \alpha)^3 = abc(1 + 3\alpha + 3\alpha^2 + \alpha^3)$. The coefficient α is always a small number, and hence α^2 and α^3 are so negligibly small that they can be discarded without producing any measurable error. Our new volume is thus very nearly $abc(1 + 3\alpha)$. The original volume abc has been increased by $3\alpha \times abc$, and the fractional increase is 3α.

The volume coefficient of a solid is, therefore, almost exactly three times its linear coefficient:

$$\beta = 3\alpha.$$

What happens to the hole in a washer when the washer is heated? Fig. 10-8A shows an aluminum washer at a temperature of 0°C. When the washer is heated to a temperature of 200°C, we realize that the width (measured as either AB or CD) must increase, and our first inclination is to assume that this increase in width will do two things: it will make the outside diameter (AD) larger; and it will also expand inward to make the diameter of the hole (BC) smaller. But before we come too firmly to this conclusion, we should look at the situation in a little more detail. The width will certainly increase, and at 200°C the increase will be

$$25 \times 10^{-6} \times 1.000 \times 200 = 0.005 \text{ in.},$$

thus making the expanded width 1.005 in. (Fig. 10-8B). As far as the

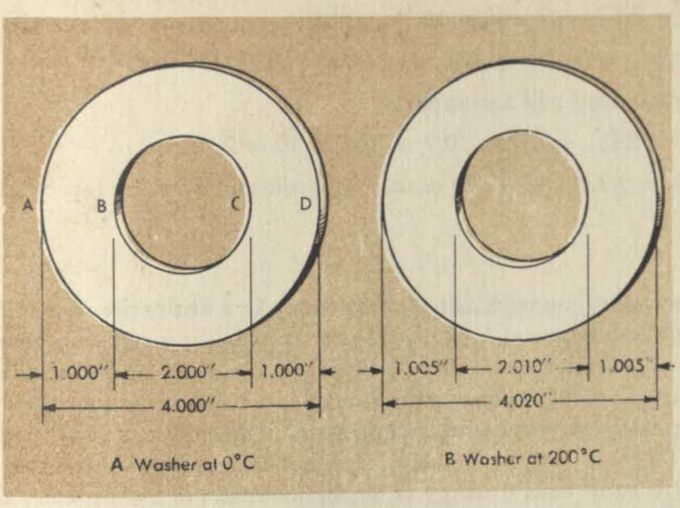


FIG. 10-8 Expansion of a heated aluminum washer.

outside diameter is concerned, it makes no difference whether the washer has a large hole, a small hole, or no hole at all; its increase will be

$$25 \times 10^{-6} \times 4.000 \times 200 = 0.020 \text{ in.},$$

to give an expanded diameter of 4.020 in. To find the new hole diameter, we need only subtract the two widths from the outside diameter, to get $4.020 - 2(1.005) = 2.010$ in. for the expanded diameter of the hole. A little checking will confirm that the hole in the aluminum washer has expanded *exactly as much as though the hole itself were made of aluminum.*

We may argue similarly that a hole inside a block of steel or glass (which is one way of describing a tank or flask) will have the same volume coefficient of expansion as the material surrounding it.

A linear coefficient of expansion for a liquid would be meaningless, but liquids do experience definite changes in volume with changes in temperature. The following list gives the volume coefficients of expansion for some common liquids at ordinary temperatures:

Ethyl alcohol	$\beta = 1.12 \times 10^{-3}/°C$
Benzene	1.24×10^{-3}
Glycerine	0.50×10^{-3}
Mercury	0.18×10^{-3}
Water	0.21×10^{-3}

Suppose a glass flask holds exactly 200 cm³ of mercury, filled to the top, at 20°C. What will happen if we heat the flask full of mercury to 80°C? The volume of the flask will increase by an amount

$$\Delta V_F = \beta \times V \times \Delta t$$
$$= 3 \times 9 \times 10^{-6} \times 200 \times 60 = 0.32 \text{ cm}^3.$$

The mercury itself will also expand:

$$\Delta V_M = 0.18 \times 10^{-3} \times 200 \times 60 = 2.16 \text{ cm}^3.$$

Thus $2.16 - 0.32 = 1.84 \text{ cm}^3$ of mercury will spill over the top of the flask.

10-7 Calorimetry

Even in everyday, nonscientific speech, we make a distinction between "heat" and "temperature," although to many people it is not very clear just what this distinction is. We "add heat" to a body and "raise its temperature," which implies, quite correctly, that heat is a quantity of something, whereas temperature is a property of the body not concerned with how large the body is. We know now that heat is a form of energy—actually, the total kinetic energy of all the molecules of a substance, as we shall see in more detail later. Temperature is a measure of the average kinetic energy of one molecule.

Heat, being energy, can be measured in foot-pounds or joules or ergs. However, it has been found convenient to define another special unit to measure heat-energy. This special unit is the *calorie* (*cal*), *which is the amount of heat necessary to raise the temperature of one gram of water by one centigrade degree.*

If we take a glass containing, say, 500 gm of water at 80°C and mix it with an equal amount of water at 50°C, we shall find that the temperature of the mixture will be 65°C, i.e., just halfway between. We have had 500 gm of water warming up by 15° from 50° to 65°, which required $500 \times 15 = 7500$ cal. These calories were furnished by the other 500 gm of water, which, in cooling 15° from 80° to 65°, gave up 7500 cal. The principle of conservation of energy works as well with calories as it does with foot-pounds or ergs.

Different substances require different amounts of heat to raise their temperatures—that is, they have different specific heats. *The specific heat of a substance is equal to the number of calories required to raise the temperature of one gram of the substance by one centigrade degree.* Most substances need considerably less than one calorie to do this. Table 10-1 lists the specific heats of a few common materials.

The specific heat of water is, of course, 1.00, because of the way in which the calorie was defined.

Some thought will enable us to write out an equation that will relate the amount of heat (Q cal), the mass of the substance (m gm), the specific heat of the substance (S cal/gm-°C), and the temperature change of the substance (ΔT°C). If we take a substance with specific heat S, we know

TABLE 10-1 SPECIFIC HEATS OF SEVERAL COMMON SUBSTANCES

Substance	Specific Heat in cal/gm-°C
Alcohol	0.52
Aluminum	0.22
Copper	0.093
Ice	0.55
Iron	0.11
Lead	0.030
Mercury	0.033

from the definition of specific heat that it will require exactly S cal to raise the temperature of 1 gm of the substance by 1°C. So for this special case of 1 gm and 1°C, we have $Q = S$. If we have m gm, it will take just m times as many calories to change the temperature by 1°C, which we can express by $Q = mS$. If the temperature is to be changed by $\Delta T°$, it will require ΔT times as many calories as it did for 1°C, and we can write as a useful relationship the following:

$$Q = mS\Delta T.$$

We should probably mention that for much of the heating and air-conditioning work in Britain and the United States, engineers use a heat unit considerably larger than the calorie. It is a British thermal unit (Btu), which is the amount of heat needed to raise the temperature of 1 lb of water by 1°F. Since this defines the specific heat of water to be 1 in these units, it is apparent that for any substance, S in cal/gm-°C $= S$ in Btu/lb-°F.

The unit used by dieticians in figuring the heat or energy content of foods is generally spelled Calorie, with a capital letter, and is equal to a thousand calories. The Calorie is also known as the *kilocalorie* (kcal).

By using the idea that the heat gained by one substance in a mixture must equal the heat lost by another, we can readily solve many problems in *calorimetry*. (Calorimetry means, literally, measuring quantities of heat.) If, for instance, we pour 400 gm of mercury at 100°C into 300 gm of alcohol at 20°C (in a cup that is insulated so that no heat can enter or leave our mixture from the outside), what will be the final temperature of the mixture? Let us call the final temperature T; T will obviously be somewhere between 100° and 20°. The mercury will cool from 100° to $T°$, a change of $(100 - T)$ degrees. It takes 0.033 cal (the S.H. of mercury) to change the temperature of 1 gm of mercury 1 degree; to change 400 gm by $(100 - T)$ degrees requires that the mercury lose $0.033 \times 400 \times (100 - T)$ cal. Similarly, to heat 300 gm of

alcohol from 20° up to $T°$ requires $0.58 \times 300 \times (T - 20)$ cal. If we set the heat lost equal to the heat gained, we get

$$
\begin{aligned}
0.033 \times 400(100 - T) &= 0.58 \times 300(T - 20) \\
1320 - 13.20T &= 174T - 3480 \\
187.2T &= 4800 \\
T &= 25.6°\text{C}.
\end{aligned}
$$

10-8 Latent Heat

When we place a teakettle on the fire, the temperature of the water gradually rises to 100°C, at which point the water begins to boil. But, once the boiling has started, the temperature stays at 100°C until the last drops of water are turned into steam. Although the heat is still flowing into the kettle from the flame, it does not make the water any hotter. What happens to that heat? The answer is, of course, that this heat is used to transform the water into vapor, and measurements show that to do it we must supply 539 calories for each gram of water to be vaporized. This amount of heat is known as the *latent* (*hidden*) *heat of vaporization* and is, of course, different for different substances. Thus to evaporate 1 gm of alcohol and 1 gm of mercury we need only 204 cal and 72 cal, respectively. The heat absorbed in the evaporation of water plays an important role during hot weather in the cooling of the body through the process of skin perspiration. Indeed, one glass of water evaporated from the surface of the body removes enough heat to cool the entire body by several degrees. Meteorologists use this principle for measuring the relative humidity of air. The apparatus used for this purpose consists of two identical thermometers with the bulb of one of them covered by a wet cloth. This thermometer, because of evaporation, shows a somewhat lower temperature, and from the difference between the two readings the weatherman can calculate the rate of evaporation and, consequently, the amount of humidity present in the atmosphere.

A similar phenomenon is encountered when water turns into ice. When the temperature of water comes down to 0°C and the first crystals of ice begin to form, the temperature remains at 0° until all the water freezes. The *heat of fusion of water* (i.e., the amount of heat that must be taken away from water at 0°C to freeze it, or be given to ice at 0°C to melt it), amounts to 80 cal/gm. The heat of fusion of alcohol (which freezes at −114°C) is 30 cal/gm, whereas for mercury (freezing at −39°C), it is only 2.8 cal/gm. To melt lead (at +327°C), it takes about 6 cal/gm, whereas in the case of copper (at +1083°C), the figure is as high as 42 cal/gm.

The idea of latent heats can add a certain amount of complication, as well as interest, to calorimetry problems. Suppose we have a calori-

meter (an insulated vessel made for doing just such an experiment as follows) containing 500 gm of water at 0°C. Into the calorimeter we put 150 gm of ice at −20°C and then bubble steam into it until the ice is melted and the entire contents are at 50°C. How many gm of steam will it take to do this? The steam passed through at first will, of course, initially condense into water, then cool to 0°, and finally warm again to 50°. However, it will give off just as much heat in cooling from 50° to 0° as it will take up later in going from 0° back to 50°. Thus we can jump directly from the beginning to the end and say that each gram of steam will give up first 539 cal in condensing into water at 100°C, then 50 cal more to cool to its final 50° temperature—a total of 589 cal, or $589x$ cal for x gm of steam. We can use this released heat energy first to warm the ice to 0°C: $0.55 \times 150 \times 20 = 1650$ cal. The ice must next be melted, which requires $150 \times 80 = 12{,}000$ cal, and gives us $150 + 500 = 650$ gm of 0° water in the calorimeter. To heat this to 50° will take another $1 \times 650 \times 50 = 32{,}500$ cal. Altogether, then, the x gm of steam must contribute $1650 + 12{,}000 + 32{,}500 = 46{,}150$ cal. Thus

$$589x = 46{,}150$$

or

$$x = 78.4 \text{ gm of steam.}$$

10-9 Heat Conduction

If we hold one end of an iron rod in our fingers and put the other end in a candle or bunsen burner flame, heat from the flame will flow through the iron into the air around the rod, and also, as we shall notice in a few minutes, into our fingers. Presently, the end we are holding will get so hot we must drop it. If this experiment is repeated with a copper rod, we must drop it much more quickly than we did the iron. We shall still, however, be able to hold a glass rod quite comfortably while the other end is red hot and molten. Apparently, some materials conduct heat much more readily than others; copper is a much better heat conductor than glass.

The rate at which heat energy flows through a rod or slab of material depends on several factors. One obvious factor is the nature of the material, and we can use a standard physicist's scheme to describe the heat conductivity of any material in a quantitative way. Imagine a cube of the material 1 cm on a side (Fig. 10-9A), with one face of the cube kept just 1°C cooler than the opposite face. Heat will flow from the warmer face to the cooler, and the number of calories per second flowing through the cube is the *heat conductivity* or *thermal conductivity* of the material. It is a simple matter to extend this concept to a piece of the material of any size and for any temperature difference. Let us figure a general formula to use on the slab shown in Fig. 10-9B, which is made of a

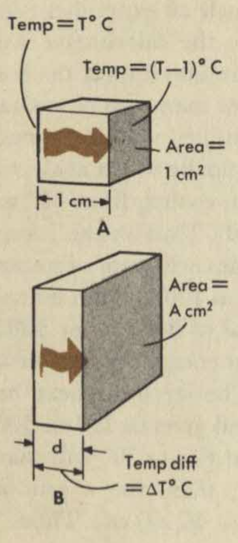


FIG. 10-9 Heat conduction through a slab of material.

material with a thermal conductivity of k. This means that *if* the slab were 1 cm thick, and *if* it had an area of 1 cm^2, and *if* one side were 1°C warmer than the other, then k cal/sec would flow through it. But it has an area of A cm^2, which considered alone would let A times as much heat flow through. The temperature difference is not 1°C, but ΔT°C, so ΔT times as much heat will flow through because of the greater temperature difference.

The thickness, d cm, will work the other way. Obviously, it is more difficult for heat to flow through a thick slab than a thin one. Experiment, as well as theoretical analysis, shows that the heat flow through a slab of material (other things being equal) varies *inversely* with its thickness. Thus since our slab is d cm thick, only $1/d$ as much heat will flow through it as through the 1-cm thickness.

The above arguments can all be assembled into a compact equation to give us the rate of heat flow Q/t:

$$Q/t(\text{cal/sec}) = \frac{kA\Delta T}{d}.$$

The heat conductivities of some familiar materials are shown in Table 10-2.

TABLE 10-2 HEAT CONDUCTIVITY OF DIFFERENT MATERIALS EXPRESSED IN CAL/SEC FLOWING THROUGH 1 CM², WHEN THE TEMPERATURE GRADIENT IS 1°C/CM

Material	Heat Conductivity (at 18°C)
Silver	0.97
Copper	0.92
Aluminum	0.48
Iron (cast)	0.11
Lead	0.08
Mercury	0.016
Glass	0.0025
Brick	0.0015
Water	0.0013
Wood	0.0003
Asbestos	0.0002
Cotton wool	0.00004
Air	0.00006

As an example, consider a jar 10 cm in diameter and 20 cm high, made of glass 2 mm thick. The lid of the jar is of copper 1 mm thick, and the water in the jar has been accidentally frozen into solid ice. The jar is put into a tank of 20°C water to thaw out. How long will it take?

First, we can find how many calories it will require to melt the ice. The volume of a cylinder is $\pi r^2 h$, or $\pi \times 5^2 \times 20 = 1570$ cm^3. Since the density of ice is about 0.92 gm/cm^3, this means there are 1450 gm of ice to be melted, which will require $1450 \times 80 = 116{,}000$ cal. The area of glass through which the heat will be conducted is $\pi \times 5^2 = 78$ cm^2 (bottom) plus $\pi \times 10 \times 20 = 628$ cm^2 (sides), for a total of $78 + 628 = 706$ cm^2, all 0.2 cm thick. The heat flow through the glass will be

$$\frac{Q}{t} = \frac{0.0025 \times 706 \times 20}{0.2} = 177 \text{cal/sec.}$$

Through the copper lid we shall have

$$\frac{Q}{t} = \frac{0.92 \times 78 \times 20}{0.1} = 14{,}300 \text{ cal/sec.}$$

This gives a total heat flow into the jar of about 14,500 cal/sec, and the required time will be $116{,}000/14{,}300 = 8.1$ sec.

Our procedure in working this problem has been perfect; the thermal conductivity values have come from dependable sources which confirm one another, and there are no large arithmetical errors. Yet the answer is ridiculously incorrect; actually, it would probably require an hour or so to melt the ice in the jar. Where have we gone wrong? The 14,300 cal/sec rate of heat transfer we computed *is* correct for the instant we put the jar into the bath; the conditions we assumed in working the problem are at this moment correct. However, in a fraction of a second the outside of the jar becomes covered with a layer of cold water; on the inside the ice is further insulated by the layer of water that has formed from its melting. Our initial conditions no longer hold, and as time goes on and more ice melts, our assumptions become more and more wrong!

A similar condition fortunately holds for windowpanes. If we calculated the loss of heat through glass windows, assuming that the inside of the glass was at a warm 70°F and the outside surface at the 0°F of the winter night, we would be shocked to discover that our entire salary could not pay our heating bill. Actually, there will be less than a 1° difference between the inside and outside surfaces of the glass; nearly all the insulation comes from a blanket of cool air against the glass on the inside, and of warm air on the outside. There are empirical rule-of-thumb multipliers experimentally worked out for use in more complicated procedures that will give approximately correct answers for the conditions common in heating design work. Pure thermal conductivity through walls and partitions is difficult to achieve except by special laboratory procedures.

Nevertheless, although we may not be able to determine the heat transfer exactly, the heat conductivity of various materials plays an important role in all kinds of heat insulation. Since cotton wool and

similar materials present 40 times more resistance to the flow of heat than ordinary brick, we can clearly see the advantages of their use for insulating homes. And, since the escape of heat is proportional to the surface of an object, to conserve heat it is advantageous to build houses as compactly as possible—hence the difference in construction styles in southern California and northern Canada. Following the same principle, many animals roll up into a ball when it is cold and stretch out when it is warm.

10-10 Heat Convection

In the case of poor heat conductors, the propagation of heat into the heated body is very slow. For example, it would take hours to heat the water in a teakettle standing on the fire if there were no other heat-carrying processes. In fluids, the propagation of heat is considerably accelerated by the process of convection, which has its basis in the fact that heated bodies increase their volume and hence decrease their density. In our teakettle, the water near the bottom is heated by immediate contact with the hot metal, becomes lighter than the rest of the water in the kettle, and floats up, its place being taken by the cooler water from

FIG. 10-10 Circulation of water in a teakettle caused by convection currents.

the upper layers (Fig. 10-10). These *convection* currents carry the heat up "bodily," and they mix the water in the kettle so that the tea is ready in almost no time. A similar phenomenon takes place in the atmosphere when, on a hot summer day, air heated by contact with the ground streams up to be replaced by cooler air masses from above. As the air rises to higher and cooler layers of the atmosphere, the water vapor in the air condenses into a multitude of tiny water droplets and forms the cumulus clouds so characteristic of hot summer days. Convection processes are also very important in the life of our sun and stars. In them, the nuclear energy produced in the hot central regions is carried toward the surface by streams of heated stellar gases.

Sometimes the notion of "convective" heat transfer gets confused with the notion of heat conduction. We have seen from Table 10-2, for example, that the heat conductivity of cotton wool is about the same as that of air. Wool, fur, and other materials used to make warm clothes also have about the same degree of conductivity. But if the heat conductivity of air is the same as that of warm clothing materials, why is a naked man less comfortable in cold weather than a man in a fur coat or under a thick woolen blanket? The reason is that the heat is removed from the skin of a naked man not so much by heat conduction into the air as by heat convection: the air warmed by contact with the skin rises and is replaced by more of the cold air. The role of warm clothing materials is to prevent this circulation, to keep the air from moving by trapping it between the numerous interwoven fibers of the materials. If we compress a woolen sweater or a mink coat under a hydraulic press, they will immediately lose much of their ability to keep us warm.

10-11 Heat Radiation

A third enormously important way in which heat energy can be transferred from one body to another is by *radiation.* If you stand outdoors warming yourself at an open fire on a winter day, the heat you receive certainly does not come to you by conduction through the air or the ground, since both of these are cold, and heat flows from you to them rather than the reverse. Neither are you heated by convection, since the hot air over the fire rises into the sky, taking its heat away with it. Just as the bright flames and the glowing coals radiate light, they also send out an even greater amount of radiant heat that travels unimpeded through the air, to be absorbed by your skin and clothing. All the energy we receive from the sun has been radiated in this way across 93 million miles of vacuum. Only a small part of this energy is in the form of light; most of the rest is radiant heat. Light and radiant heat are two forms of electromagnetic radiation, which we shall investigate later in much greater detail.

10-12 Very Hot and Very Cold

We human beings are very sensitive to temperature. If it is in the "upper 80s" (Fahrenheit), we feel uncomfortably hot; and if it is "below freezing," we feel cold. Actually, life can survive much wider temperature variations. Algae living in hot springs can stand temperatures almost as high as the boiling point of water, and on the other hand some plant seeds are not killed by temperatures approaching absolute zero. The lowest temperature we encounter in everyday life is probably that of dry ice (frozen carbon dioxide), which is about −80°C, and the highest is that of the kitchen range flame, which is about +1700°C. High-temperature engineering goes well beyond this in the production of high temperatures; a "plasma-jet" torch, for instance, reaches a temperature of almost 15,000°C. All solid bodies melt and then evaporate when the temperature is sufficiently high, and in Table 10-3 we list some melting and boiling points to give an idea of their range of temperatures. At temperatures of 6000°C and above, all (even the most heat-resistant) materials are turned into gas, which is the state of all materials in the atmosphere of our sun.

TABLE 10.3 MELTING POINTS AND BOILING POINTS

Substance	Melting point (°C)	Boiling point (°C)
Helium	−272	−269
Hydrogen	−259	−253
Nitrogen	−210	−196
Oxygen	−218	−183
Chlorine	−101	−35
Water	0	+100
Tin	+232	+2260
Lead	+327	+1620
Aluminum	+660	+1800
Copper	+1083	+2300
Iron	+1535	+3000
Platinum	+1773	+4300
Tungsten	+3370	+5900

Questions

(10-1)

1. The average normal temperature of the human body is taken to be 98.6°F. What is this temperature on the Centrigrade scale?

2. The melting point of the element sulfur is 113°C. What is this temperature on the Fahrenheit scale?

3. The Reaumur temperature scale is almost obsolete, but is still used to some extent in France. On this scale, the freezing point of water is 0°, and the boiling point 80°. What Centigrade and Fahrenheit temperatures correspond to 32°R?

4. Normal room temperature (70°F) corresponds to what temperatures on the Centigrade and Reaumur scales? (See Question 3.)

5. A motor is designed to heat up to 50°C above room temperature when running continuously at full load. What would its temperature be in °F when it is operating in a room at 80°F?

6. A motor is designed to heat up to 50°C above room temperature when running continuously at full load. In how hot a room (in °F) could it be operated if its temperature is not to exceed 200°F?

(10-3) **7.** Convert the following temperatures to absolute temperatures (°K): (a) 120°C, (b) 1500°F, (c) −30°C, (d) −78°F.

8. Convert the following temperatures to absolute temperatures (°K): (a) 350°C, (b) 120°F, (c) −250°C, (d) −130°F.

(10-4) **9.** What is a standard atmosphere of pressure in lb/in^2?

10. On a mountain, the barometric pressure is 60 cm Hg. Convert this pressure into (a) dynes/cm^2, (b) lb/in^2. (See Question 9.)

11. A compressor takes 100 ft^3 of air at standard atmospheric pressure and forces it into a tank whose volume is 8 ft^3. What is the pressure of the gas in the tank? (Assume temperature unchanged.)

12. A tire holds 1200 in^3. What volume of air at standard atmospheric pressure would have to put into the tire to inflate it to a pressure of 40 lb/in^2?

(10-5) **13.** A bottle is tightly sealed at a temperature of 70°F, and is then put into a 350°F oven. (a) What is the absolute pressure of the air in the bottle at this higher temperature? (b) What is the gauge pressure? (All this takes place at sea level on an average day.)

14. A bottle is tightly sealed at a temperature of 18°C, and is then placed in a 210°C oven. The atmospheric pressure is 13.0 lb/in^2. (a) What is the absolute pressure in the bottle at this high temperature? (b) What is the gauge pressure?

15. A man checks his tire gauge pressure as 28 lb/in^2 in the early morning when the temperature is 41°F. In the afternoon, hard driving on a hot pavement has raised the temperature of his tires to 140°F. (a) What is the gauge pressure of his tires, if the atmospheric pressure has not changed? (b) Does it make any difference in your answer whether atmospheric pressure is taken to be 10 lb/in^2 or 15 lb/in^2?

16. A compressed-air tank in a filling station has a gauge pressure of 50 lb/in^2 when the temperature is 70°F. A fire in the station raises the temperature of the tank to 250°F. What is the gauge pressure in the tank when hot? (Atm. pr. = 15 lb/in^2.)

17. A large plastic balloon will hold 10,000 ft^3 when completely filled. It is designed to rise to an altitude of about 10 km, where the pressure will be 22 cm Hg, and the temperature −70°F. How many cubic feet of helium gas at standard atmospheric pressure and 70°F should be put into the balloon if it is to be full when it reaches its designed altitude?

18. A tank is guaranteed safe at 100 lb/in^2 gauge pressure. It is filled with gas to 80 lb/in^2 gauge pressure at sea level and 20°C. If it is taken up in a plane to where atmospheric pressure is 5 lb/in^2, how hot may the tank be safely allowed to get? (Take sea level atm. pr. = 15 lb/in^2.)

(10-6) **19.** A steel surveyor's tape is exactly 300 ft long under a certain specified tension at 20°C. How long is the tape when the temperature is −10°C?

20. A steel bridge is exactly 250 ft long at 15°C. How long is the bridge when its temperature is 42°C?

21. A steel ball 9 cm in diameter has a diameter 0.012 cm too large to permit it to fit into a hole in a brass plate when the temperature is 20°C. (a) Should the ball *and* plate be heated or cooled in order for the ball to fit the hole? (b) At what temperature will the ball fit the hole?

22. At 15°C the diameter of an aluminum rod is 0.0005 in. too large to fit into an exact 1-in. diameter hole in a brass plate. (a) Should the rod *and* plate be heated or cooled in order to make a fit possible? (b) At what temperature will the rod fit in the hole?

23. A block of aluminum is exactly 10 cm × 10 cm × 10 cm at 0°C. What is the volume of the block when it is heated to 80°C?

24. A glass flask holds exactly 200 cm^3 at 15°C. What is its capacity when heated to 60°C?

25. A vertical glass tube with the bottom end sealed is filled with mercury to a height of 100 cm, at 15°C. How high will the mercury stand in the tube if the temperature is raised to 35°C?

26. A 2000-cm^3 aluminum tank is filled with water at 80°C. How much more water can be added when the tank and contents have cooled down to 10°C?

(10-7) **27.** How much heat is required to raise the temperature of 700 gm of mercury from 50° to 122°F?

28. An insulated tank car holds 15,000 kg of alcohol at 25°C. How much heat must be removed from the alcohol to cool it down to 12°C?

29. How much heat must be added to 50 kg of ice at −40°C to raise its temperature to −5°C?

30. How much copper shot at 95°C must be added to 150 gm of water at 25°C in order to heat it up to 39°C?

31. If 300 gm of iron shot at 5°C are poured into an insulated cup (called a calorimeter) which contains 500 gm of alcohol at 70°C, what is the final temperature of the mixture?

32. 500 gm of powdered aluminum at 100°C is mixed with 200 gm of iron filings at 20°C in an insulated cup (calorimeter). What is the final temperature of the mixture?

33. In Question 31, we have not taken into account the heat needed to change the temperature of the calorimeter itself. Suppose it is made of copper, with a mass of 80 gm. Refigure Question 31 using this additional information. (The calorimeter is, of course, always at the same temperature as its contents.)

34. In Question 32, we have not taken into account the heat needed to change the temperature of the calorimeter. Suppose it is aluminum, with a mass of 35 gm. Refigure Question 32, using this additional information.

(10-8) **35.** A cup contains 300 gm of too-hot coffee at 95°C. (a) How much water at 0°C must be added to reduce the temperature to 75°C? (b) How much ice at 0°C must be added to reduce the temperature to 75°C?

36. An insulated tank holds 500 kg of water at 20°C. (a) How much water at 100°C must be added to raise the temperature to 45°? (b) How much steam at 100°C must be added to raise the temperature to 45°C?

37. Suppose 200 gm of crushed ice and 600 gm of iron shot, all at 0°C, are in a 100-gm aluminum calorimeter. If 50 gm of steam at 100°C are slowly bubbled through, what is the final temperature?

38. 400 gm of crushed ice and 200 gm alcohol are at −25°C in a 75-gm copper calorimeter. If 120 gm of steam at 100°C is bubbled through this mixture, what is the final temperature?

(10-9) 39. If one end of a thermally insulated copper rod 1 m long and 2 cm in diameter is placed in boiling water and the other end in ice water, how many calories will be transmitted along the rod in one minute?

40. An aluminum rod 40 cm long has a cross section area of 5 cm^2, and is thermally insulated. One end is kept at −78°C in a bath of dry ice and acetone; the other end is in an insulated container of 100 gm of water at 10°C. How long will it take to cool the water down to 0°C?

41. A picnic refrigerator is a box 20 cm × 40 cm × 50 cm made of wood 5 cm thick. It contains 5 kg of ice. If we assume the interior temperature stays constant at 41°F while the outside temperature is 95°F, how long will it take for the ice to melt? (There is a drain hole, through which the water leaks out as the ice melts.)

42. An insulated vat is divided into two sections by a watertight partition 20 cm × 40 cm made of copper 1 mm thick. On one side is water which is boiling, and on the other side a mixture of 2 kg of crushed ice kept constantly stirred in water. How long will be required to melt all the ice?

43. Take two houses of equal volume, one 20 × 90 ft × 8 ft high and the other built in three stories 20 ft × 30 ft × 24 ft high. Approximately how will the heating bills of the two houses compare, if insulation, window area, weather, etc., are equal?

44. Consider two greenhouses of equal height and the same floor space: one is 15 ft × 120 ft, the other 45 ft × 40 ft. How will their heating bills compare?

45. When we step barefooted out of bed on a winter morning, it is much more comfortable to step on a rug than a wooden floor, although both are at the same temperature. Explain.

46. Numerous small children have their tongues severely injured by licking a piece of pipe or other metal on a very cold day; licking a piece of wood on the same day would result in no harm. Explain.

chapter / eleven

Heat and Energy

11-1 Mechanical Equivalent of Heat

In an earlier chapter of this book, we saw that friction forces gradually rob mechanical systems of their energy and eventually bring them to a standstill. What is the relation between the mechanical energy lost to friction and the amount of heat produced by it? This question was answered in the middle of the last century by a British physicist, James P. Joule (1818–1889), in his famous experiment on the transformation of mechanical energy into heat. Joule's apparatus, schematically shown in Fig. 11-1, consisted of a water-filled vessel containing a rotating axle with several stirring paddles attached to it. The water in the vessel was prevented from rotating along with the paddles by special vanes attached to the walls of the vessel. The axle with the paddles was driven by a weight hanging from a cord, and thus the work done by the descending weight was transformed by friction into heat produced in the water. Knowing the amount of water in the vessel, Joule could measure the rise of its temperature and calculate the total amount of heat produced; the driving weight and the distance of its descent gave the total amount of mechanical work done. Repeating this experiment many

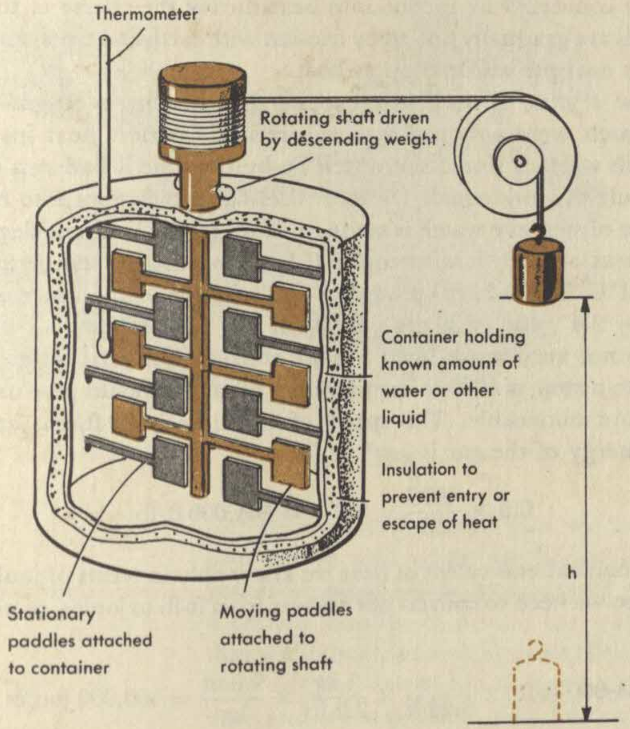


FIG. 11–1 Joule's apparatus for determining the mechanical equivalent of heat.

times and under different conditions, Joule established that there is a direct proportionality between these two quantities and that "the work done by the weight of 1 pound through 772 feet at Manchester will, if spent in producing heat by friction in water, raise the temperature of 1 pound of water one degree Fahrenheit." In metric units, it means that *one calorie of heat is the equivalent of 4.18×10^7 ergs of work, or 4.18 joules.*

Joule's work confirmed the basic idea that was commencing to be seriously considered at the time, namely, that ***heat is energy in the same sense that mechanical energy is, and although one form of energy can be transformed into another (kinetic, potential, heat, electrical, chemical, etc.), the total sum of the energy in any system of bodies or materials remains constant.*** (This assumes, of course, that our "system" does not receive any energy from the outside, or lose energy to it.) This law of the *conservation of energy* represents one of the basic pillars of physics and is known as the *first law of thermodynamics.*

The fate of all the mechanical energy on the earth (as well as of electrical and other forms of energy) is to be converted into heat at the exchange rate of 4.18 joules per calorie. The work of winding your watch, which is stored temporarily as the potential energy of the coiled spring, is converted into the kinetic energy of the escapement wheel and

gradually converted by friction into heat during the course of the day. The winds are gradually slowed by friction with earth and trees, and their enormous energies also end up as heat.

Suppose 1 joule of work were expended in winding a 30-gm watch. If the watch were wrapped in a theoretically perfect heat insulator, how much warmer would the watch be by the time it had run down? The 1 joule of work equals $1/4.18 = 0.24$ cal, which goes into heating the 30 gm of steel the watch is made of. If we take 0.1 cal/gm-deg as the specific heat of steel, it will require 0.1 cal to raise the temperature of 1 gm by 1°C. The 0.24 cal produced will raise the temperature of 1 gm of steel by 2.4°, and of 30 gm, just $\frac{1}{30}$ of 2.4° or 0.08°C.

This is not very much heat, but we started with very little energy. Braking to a stop, a 4000-lb car traveling 60 mi/hr should give us something more noticeable. The speed of 60 mi/hr is 88 ft/sec, and the kinetic energy of the car is $\frac{1}{2}mv^2$ or

$$0.5 \times \frac{4000}{32} \times (88)^2 = 484{,}000 \text{ ft-lb}.$$

The mechanical equivalent of heat we know only in terms of joules and calories, so we need to convert our answer from ft-lb to joules, or newton-meters:

$$484{,}000 \text{ ft-lb} \times \frac{1 \text{ m}}{3.28 \text{ ft}} \times \frac{1 \text{ kg}}{2.2 \text{ lb}} \times \frac{9.8 \text{ nt}}{\text{kg}} = 660{,}000 \text{ joules}$$

and

$$660{,}000 \text{ joules} \times \frac{1 \text{ cal}}{4.18 \text{ joules}} = 157{,}000 \text{ cal}.$$

This amount of heat is enough to raise 1570 gm of water (about 3 pints) from freezing to boiling, or to raise about 40 lb of steel from room temperature to the boiling point of water.

11-2 Heat and Mechanical Energy

We know that any amount of mechanical energy can be completely transformed into heat and thus that all the kinetic energy of a heavily loaded, fast-moving train is transformed into heat by the friction in the brakes when the train stops. But is the process reversible? Can the entire amount of heat contained, let us say, in a pot of boiling water be transformed into mechanical energy? We know, for sure, that steam engines do transform heat into mechanical energy, but if we look at the problem more closely, we shall find that they always transform only a part of the entire heat available into mechanical energy.

To understand why, let us consider the following problem, which at first sight has nothing to do with heat or steam engines. Suppose there is a house on a hill and a creek running swiftly through a ravine a dozen

FIG. 11-2 Pumping water uphill by means of a water wheel.

feet below it (Fig. 11-2). Can we manage it so that the creek by its own power will supply the household's water needs? The answer is yes. If we build a dam A and install a water wheel B that will produce a certain amount of power, the water wheel can operate a pump C that will pump a certain amount of water up the hill and into the house. Very simple indeed! But if the owners of the house become too ambitious and try to get more than that certain amount of water up the hill, they will be heading for trouble. The amount of water they are getting is being pushed 12 ft uphill by the rest of the water dropping 3 ft in the water wheel. If all the water in the creek were brought up to the house, there would be no water left to drive the wheel and to operate the pump! The best we can do is to arrange things in such a way that the potential energy liberated by the water operating the wheel is the same as the energy necessary to raise the water to the house. If x is the fraction of the total water supply of the creek that can be brought up to the house, then $(1 - x)$ is the fraction remaining that will run over the water wheel and produce power, and we have

$$x \times 12 = (1 - x) \times 3$$

so that

$$x = \frac{3}{12 + 3} = \frac{1}{5}.$$

Thus, at best, we can get one-fifth of the water of the creek self-propelled to the house, and any demand beyond that would contradict the laws of physics.

The situation with any heat engine (a steam engine or an automobile engine, for example) is quite similar. It was shown by a young French army engineer, Nicholas Sadi Carnot (1796–1832) that ***if an engine is***

operated by heat energy flowing from a high temperature T_H to a cooler temperature T_C, the greatest possible fraction of the heat energy that can be converted into mechanical work is given by $(T_H - T_C)/T_H$. (The T's, of course, are *absolute* temperatures.)

As an example, we can consider a small steam engine whose boiler temperature is 150°C, and in which the exhaust is cooled down to 40°C (423°K and 313°K, respectively, on the absolute scale). The maximum fraction of the heat from the boiler that can be transformed into mechanical work is

$$\frac{423 - 313}{423} = \frac{110}{423} = 0.26, \quad \text{or 26 percent.}$$

Any actual engine is much less efficient than Carnot's theoretically perfect engine, and an actual engine operating at these temperatures would do well to convert 15 percent of the heat supplied to the boiler.

11-3 Entropy

If we could transform 100 percent of a given amount of heat into mechanical energy, we would not have to worry about mining coal or drilling for oil and gas. The first law of thermodynamics (which is only the law of conservation of energy) does not say that it cannot be done. An ocean liner could pump in sea water, extract the heat energy from it to drive its propellers, and throw overboard the resulting chunks of ice. An airplane could take in air, turn its heat into kinetic energy, and throw an ice-cold jet out through a nozzle in the rear. But, as Carnot demonstrated, we cannot use the heat content of our surroundings to produce mechanical work any more than we can use the water of the oceans to operate hydroelectric power installations. The potential energy of the water in the oceans is useless because there is no lower water level to which their water could flow; and the heat content of our surroundings is useless because there is no lower temperature region to which this heat can flow.

The big electric generators at Niagara Falls are possible only because water can be drawn from the river above the falls, its potential energy used to drive water wheels, and then discharged at the lower level of the river below the falls. Here the transformations of energy are entirely mechanical, but the basic principle applies equally well to heat engines.

All heat-operated generator plants (whether the fuel be coal, oil, or gas, or a nuclear reactor) are located on a river or a lake in order to have a large cool reservoir (T_C) into which the used heat (originally at a temperature T_H) can be discharged.

Obviously, some sort of organization is needed for us to be able to derive useful work from a system. We must have two water reservoirs at different elevations to run a water wheel; and we must have two heat reservoirs at different temperatures to run a heat engine. If we have two insulated water tanks in a laboratory—one containing 100 gallons of hot water at 95°C, and the other 100 gallons of cold water at 5°C—

we can imagine ourselves extracting considerable work from this system by shuttling a properly designed cylinder and piston back and forth between the two tanks until they are finally at the same temperature. If, however, we mix the two tanks together, we shall have merely 200 gallons of 50°C water from which we can extract no useful work (unless, of course, we bring in some more water at a different temperature).

As far as the first law of thermodynamics is concerned, 200 gallons of 50°C water contains exactly as much heat energy as our two separate tanks at 95°C and 5°C. What is missing is the *organization*, or *orderliness*, of the two separate tanks, carefully segregated with hot water in one tank and cold water in the other.

Scientists have a very useful word—*entropy*—which measures the *dis*-orderliness or *lack* of organization in a system. If you throw a piece of red-hot metal into a can of water, you have at the beginning a neatly ordered arrangement of the available heat of the metal-plus-water system; the high-temperature material is all in one place, and in the remainder of the can, well separated from the hot metal, is the lower-temperature water. The entropy of the system is relatively low. But in a few minutes this ordered arrangement is lost. The heat energy that was separated into hot material at one place and cool material at another is now randomly and equally distributed, and metal and water are all at the same temperature. The entropy of the system has *increased*.

This is the natural way for any transfer of heat energy to take place. If two bodies or materials of different temperature are placed together so that heat can be transferred, the hot body will cool, and the cold body will warm until the final temperature is the same for both. In other words, the natural way of the world is to equalize energy differences—more specifically, differences in temperature. Entropy is a measure of how well these differences are wiped out.

Mathematically, the definition of a change in entropy is very simple:

$$\Delta S = \frac{\Delta Q}{T}$$

where S represents entropy and ΔS is the change in entropy; ΔQ represents the heat added (it is negative if heat energy is lost); and T is of course the absolute temperature at which the ΔQ was either lost or gained. We shall get into trouble if we try to apply this to the hot metal thrown into the can of water; the temperatures of the metal and of the water are constantly changing, which makes calculus necessary for handling the problem. We can, however, think of simple cases in which this difficulty does not arise.

Consider, for example, a 100-gm ice cube (0°C) thrown into a large container of warm water (40°C). In order to melt, the ice cube must absorb $100 \times 80 = 8000$ cal, and its temperature remains at 273°K until it is all melted. One ice cube will not measurably change the temperature of the water, so we can consider this, too, to be constant at

313°K. The entropy change of the ice cube is expressed as $\Delta S_i = 8000/273 = 29.3$ cal/deg. For the water (which has lost heat), the change is $\Delta S_w = -8000/313 = -25.6$ cal/deg. Thus the total entropy change in the ice-and-water system is $29.3 - 25.6 = +3.7$ cal/deg. The entropy has increased, as is always true for any spontaneous natural process, if we are careful to include everything that is involved.

The "natural" direction of the flow of heat energy is from a warmer body to a cooler body, in agreement with the rule that entropy must always increase. We can think of examples in which this seems to be violated. In our refrigerators, for example, heat is taken from cold water to freeze it into ice cubes; and this heat taken from the water is discharged into the hot kitchen, to make it still hotter. This looks like a definite *decrease* in entropy and a flow of heat in the wrong direction. And of course it is, but only because we have not included the entire extent of the system involved. In the preceding paragraph, the water experienced an entropy decrease, but this was more than offset by the increase in entropy of the ice. Your refrigerator operates only if an electric motor is busily at work inside it. The entire system involved includes not only the freezing water and the room but also the motor and the power plant that generates the electric power to make it run. If we include all of this system, the negative ΔS of the refrigerator itself is more than offset by a much larger positive ΔS in the electric power plant, so that the total ΔS is positive, as it must always be.

Questions

(11-1)

1. Confirm the statement at the end of Sec. 11-1, that 1.57×10^5 calories will heat 40 lb of steel as stated.

2. Confirm, by using Joule's quoted figures, that 1 cal $\approx$ 4.18 joules. (This will not come out exactly, as more modern work has shown that Joule's figures, based on relatively crude apparatus, were a little off.)

3. A 500-gm weight is dropped 150 cm to the floor. How many calories of heat are produced?

4. In a paddle-wheel experiment like Joule's, a 20-kg mass descends through a distance of 1.25 m. After each descent the weight is wound back up and allowed to come down again. This is repeated 24 times. The insulated container holds 2800 gm of water. What rise in the temperature of the water could be expected?

5. A 10-gm bullet traveling 300 m/sec hits a target and is quickly stopped. (a) How much KE is converted into heat? (b) How many calories is this?

6. A 1500-kg car traveling 72 km/hr is brought to a stop (whether by a stone wall or the car's brakes makes no difference). How many calories of heat are produced?

7. A 10-kg piece of lead is dropped from a building 100 m high. Assume that 0.8 of its energy is converted into heat in the lead, and the remaining 0.2 into

heat in the ground. (a) How many calories go into heating the lead? (b) How many degrees will the temperature of the lead rise?

8. A 20-kg piece of aluminum is dropped from a building 150 m high. Assume that 0.7 of its energy is converted into heat in the aluminum, and the remaining 0.3 into heat in the ground. (a) How many calories go into heating the aluminum? (b) By how many degrees will the temperature of the aluminum rise?

9. A lead bullet of unknown mass is fired at a speed of 200 m/sec into a tree, in which it stops. Assuming that $\frac{2}{3}$ of the heat produced goes into the bullet and $\frac{1}{3}$ into the wood, find how many degrees the temperature of the bullet is raised.

10. An iron bolt is fired at a speed of 300 m/sec, by a special gun, into a concrete wall (a common procedure in modern construction practice). Assuming that $\frac{2}{3}$ of the heat produced goes into the iron, find the resulting rise in temperature of the bolt.

11. A sack of ice at 0°C is dropped from a plane flying at an altitude of 500 m. What fraction of the ice will be melted by its impact with the ground, if half of the heat produced is absorbed by the ice?

12. The melting point of lead is 327°C. At what speed would a lead bullet have to be fired into a target in order to completely melt the bullet, if half of the heat produced is absorbed by the bullet? (Bullet initially at 27°C.)

13. A 100-watt lamp is immersed in an insulated container holding 500 gm of alcohol at 20°C. How long will it take to warm the alcohol to 50°C?

14. A 1500-watt heating coil is immersed in an insulated tank holding 100 kg of ice at −10°C. How long will it take to raise the temperature of the contents to 70°C?

(11-2) **15.** What is the maximum possible thermal efficiency (i.e., the fraction of energy supplied which is converted into useful work) that can be obtained from an engine whose heat input is at 400°C and which discharges its waste heat at 60°C?

16. What is the maximum possible thermal efficiency of an engine whose heat input is at 518°F and which discharges its waste heat at 176°F?

17. An engine operates between the temperatures 800° and 200°C. By what fraction would its theoretical thermal efficiency be increased if a redesign lowered its discharge temperature to 100°C?

18. An engine operates between the temperatures 800° and 200°C. By what fraction would its theoretical thermal efficiency be increased if it were found possible to raise its high temperature to 900°C?

(11-3) **19.** A heating coil is immersed in an insulated tank that contains 1000 gm of crushed ice mixed with an equal mass of water, all at 0°C. Steam at 100°C is passed through the coil until all the ice is just melted. (a) What has been the entropy change of the steam? (Assume it leaves the far end of the coil as 100°C water.) (b) What has been the entropy change of the ice? (c) What has been the total entropy change?

20. A tank is divided into two compartments by a partition of sheet copper. On one side is water at 80°C; on the other is water at 20°C. In a brief period of time, 1000 calories will flow from the hot side to the cool. What is the total change in entropy during this period? (Assume there is so much water that 1000 calories does not change the temperature appreciably.)

chapter / twelve

Electrostatics

12-1 Static Electricity

In the early experiments with electricity carried out by William Gilbert (1540–1603), a personal physician of Queen Elizabeth I, electric charges were produced mainly by "rubbing a galosh against a fur coat" or a glass rod against a silk handkerchief. When a lady with a fur coat gets into a car with plastic seatcovers, or a gentleman wearing rubber-soled shoes walks across a carpet, sparks may fly when they touch the handle of the car door, or the radiator in the room. Electricity produced in this and similar ways is known as *static electricity*, and from early studies of it the first laws of electric interactions were published.

If we suspend, side by side, two light metallic or metal-covered balls and touch each of them with a stick of hard rubber that has been rubbed with a piece of fur or wool, we shall find that the balls repel each other (Fig. 12-1A). Repulsion also takes place if we touch both balls with a piece of fur or wool that has been rubbed against hard rubber (Fig. 12-1B). However, if one sphere is touched by the hard rubber stick and the other by the fur, the two spheres will attract each other (Fig. 12-1C).

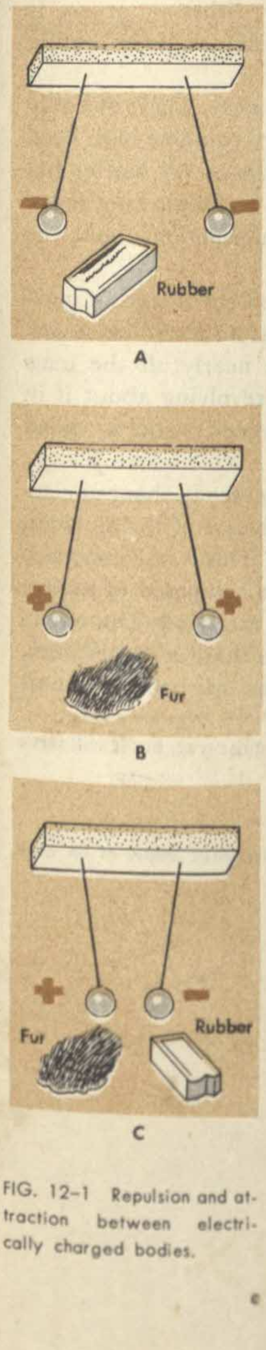


FIG. 12-1 Repulsion and attraction between electrically charged bodies.

On the basis of such elementary experiments, Gilbert concluded that there are two kinds of electricity and that ***electric charges of the same kind repel each other, while those of the opposite kind attract each other.***

In the late eighteenth century, the American scientist and statesman Benjamin Franklin named the electric charges produced on the fur by friction *positive* and those produced on the rubber *negative*. Studying in more detail the interactions between electric charges, the French physicist C. A. Coulomb (1736–1806) found that ***the force of attraction or repulsion between two electrically charged bodies varies in direct proportion to the product of their charges* and in inverse proportion to the square of the distance between them.*** *Coulomb's law* formed the cornerstone for all further studies in the field of electricity. Following the general method of defining the units for new physical quantities, *we can define the electrostatic unit of charge (esu) as the amount of electricity which acts with a force of one dyne on an equal amount placed one centimeter away.* Strictly speaking, we should add to the above definition that the charges must be in a vacuum. However, the difference is very small if they are in air.

With our unit of charge thus defined, we can write out Coulomb's law in mathematical form, exactly as we did previously for Newton's very similar law of gravitation:

$$F = \frac{Q_1 Q_2}{d^2}.$$

From the way in which the esu was defined, we can see that in this statement of Coulomb's law, F must be in dynes, Q in esu, and d in cm. The esu is too small for practical purposes, and in electrical engineering we use a much larger unit known as a *coulomb* (coul), which is equal to approximately 3 billion (3×10^9) electrostatic units. We shall find occasion to make use of the coulomb a little farther along.

As an example of Coulomb's law, let us calculate the force of repulsion between two electrons 5×10^{-8} cm apart. Electrons are tiny negatively charged particles, each having a charge of 1.60×10^{-19} coul. To use Coulomb's law in the form given here, the charges must be in esu:

$$1.60 \times 10^{-19} \text{ coul} \times \frac{3 \times 10^9 \text{ esu}}{1 \text{ coul}} = 4.80 \times 10^{-10} \text{ esu}.$$

Then

$$F = \frac{(4.80 \times 10^{-10})^2}{(5 \times 10^{-8})^2} = 9.22 \times 10^{-5} \text{ dyne}.$$

* Coulomb obtained half-unit, quarter-unit, and smaller integral fractions of charges for his experiments by taking a charged sphere and bringing it in contact with uncharged ones of equal size, the charge being equally divided between the balls in contact.

12-2 Elementary Atomic Structure

For an understanding of how an electric charge, either + or −, is associated with material bodies, we must take a brief look at the atomic structure of matter.

All matter is made up of tiny particles known as *atoms*. There are only about a hundred different kinds of atoms, and they combine with each other in different ways to form groups called *molecules*. All matter has been found to be composed of atoms or molecules, and some knowledge of how atoms are made will give us valuable information about the behavior of matter.

In 1911, Lord Rutherford in England discovered (by means of ingenious experiments we shall discuss farther on) that an atom has a tiny *nucleus* which is positively charged and contains nearly all the mass of the atom. Distributed about the nucleus and revolving about it in orbits* are much less massive negatively charged particles called *electrons*.

In a normal atom, there are exactly as many negatively charged electrons as are needed to neutralize the positive charge of the nucleus, so that the atom as a whole is electrically neutral. This is of course also true of all normal material substances, which are composed of atoms. Figure 12-2 gives a rough idea of how an atom is arranged. The outermost electrons are less strongly bound to the atom than the inner ones, and they are the ones that take part in chemical reactions between atoms

* The visual picture of electrons revolving in orbits is not quite true, but it will serve as a tentative model.

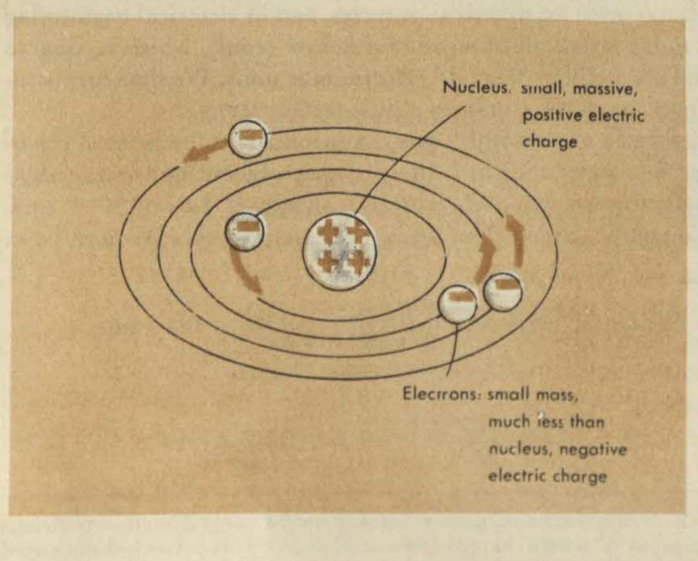


FIG. 12–2 Schematic model of an atom.

and that are responsible for the accumulation of an electric charge on bodies.

Not all atoms or groups of atoms hold their outer electrons equally firmly. The attraction, or affinity, of hard rubber for electrons is greater than that of fur fibers. Accordingly, when a hard rubber stick is rubbed with fur, some of the electrons leave the fur and become attached to the rubber. Thus the hard rubber, with an excess of negative electrons, becomes negatively charged; the fur, left without enough electrons to neutralize the positive charges of its atomic nuclei, is left with a net positive charge. If a glass rod is rubbed with a piece of silk, the glass will be found to have a positive charge and the silk a negative charge, indicating that the silk fibers have a greater affinity for electrons than the glass does. We must bear in mind that, in solid materials, only some of the outer electrons may have any freedom to move. The remainder of the atom is firmly fixed in place.

12-3 Conductors and Nonconductors

Let us take a ball made of cork or pith, covered with metal foil, and touch it with a hard rubber rod previously rubbed with fur or a woolen cloth. Some of the excess electrons on the rod will leave it and transfer to the ball, giving the ball a negative charge. The negative charge on the ball can be confirmed by noting that it is strongly repelled by the rubber rod. (Like charges repel each other.)

Now if you hold in your hand a piece of bare wire, a metal spoon, or a wet stick of wood and touch it to the charged ball, the ball will be found to lose its charge. Apparently, the excess of electrons on the ball is able to flow into your hand and body through the connecting material. If this experiment is repeated, and you hold a piece of dry wood, a glass rod, a tube of paper, or a piece of plastic, the charge will be found to remain on the ball.

This experiment can serve as a basis for separating all materials into two general classes: *conductors*, through which a stream of electrons can readily flow, and *insulators*, or *nonconductors*, through which electrons cannot move. In general, metals and water solutions of salts, acids, and alkalis are conductors, whereas nonmetallic solids and oils are nonconductors. Do not take this as a rigorous classification scheme, because all conductors offer some resistance to the passage of electrons and all insulators permit some few electrons to pass through. A number of materials are on middle ground, and are neither good conductors nor good insulators. For many purposes, however, the classification is a useful one.

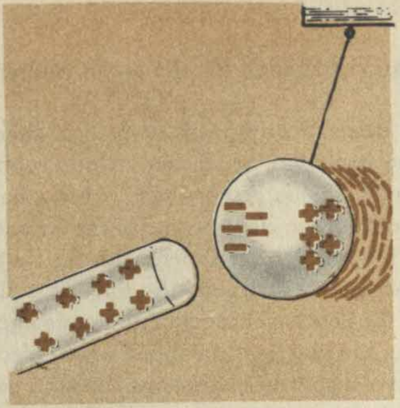


FIG. 12–3 The attraction of an electrically neutral foil-covered pith ball to a charged rod.

12-4
Induced Charges

We can be sure a foil-covered pith ball will be electrically neutral if it is touched with a finger or with one end of a wire connected to a water pipe that is in turn buried in the ground. Any excess or deficiency of electrons on the ball will be removed by a flow of electrons to or from your body. Yet such a neutral ball will be attracted by either a positively or a negatively charged rod. We must look for some explanation of this behavior, since we might expect from Coulomb's law, with one of the Q's $= 0$, that a neutral ball would not be affected one way or the other. Figure 12-3 shows what happens in such a situation. The attraction of the positively charged rod draws toward itself some of the mobile electrons in the metal covering of the ball, thus leaving an equal number of the atoms on the distant side of the ball with a net positive charge. The rod consequently attracts the near side of the ball and repels the far side; since the near side is nearer (naturally!), the attraction is stronger than the repulsion, and the net effect is a force of attraction. We need only reverse all the signs in Fig. 12-3 to see that a negatively charged rod would also, and equally, attract a neutral ball.

When the charged rod is removed from the neighborhood of the ball, the excess electrons on the front of the ball return to the rear half, and the ball is again uniformly without charge of either kind. However, if we touch the ball with a finger *while the charged rod is still near*, the situation is changed. Figure 12-4 shows electrons, attracted by both the rod and the electron-deficient half of the ball, flowing into the ball. The finger is then removed, and if the charged rod is taken away, the excess electrons, as a result of their mutual repulsion, distribute themselves evenly over the surface of the ball and give it a uniform negative charge. If the same experiment were repeated with a negative rod, electrons would be repelled from the ball into the finger and the ball would be charged positively.

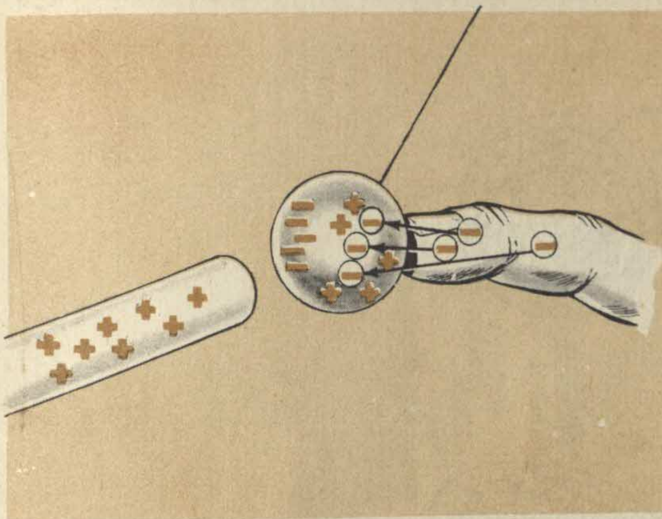


FIG. 12–4 Electrons flow from finger to ball, thus giving the ball a charge by induction.

The attraction of a neutral foil-covered pith ball toward a charged rod was easily explained in terms of the migration of electrons through the conducting foil. However, the charged rod will also attract pieces of dry paper, lint, and hair, which are all good insulators, so that the electron migration explanation will not work. Here we have an example of the *polarization* of atoms in an electric field. We have pictured an atom as consisting of a massive positive nucleus surrounded by a whirling cloud of negative electrons. Normally, the electron cloud is centered on the nucleus, but when a nearby charged rod creates an electric field, the atoms become distorted. Figure 12-5A shows a normal atom, with the electron cloud centered on the nucleus. In Fig. 12-5B, the nearby negatively charged rod attracts the nucleus and repels the electrons, with the result that the center of the electron cloud now falls on the far side of the nucleus, so that the attraction for the nucleus is greater than the repulsion of the electrons, and the atom as a whole is slightly attracted toward the rod. This effect is small for any one atom, but there are an enormous number of atoms in a piece of paper or a bit of hair, and the total force of attraction may be quite appreciable. You should redraw Fig. 12-5 for yourself, with a positive rod, to demonstrate that this, too, will result in a polarization of the atom in such a way that it will again result in an attraction.

12-5 The Electroscope

For the detection and measurement of electric charge, many different devices have been designed, some of them complicated and expensive electronic instruments. Among the simplest, however, is the *gold-leaf*

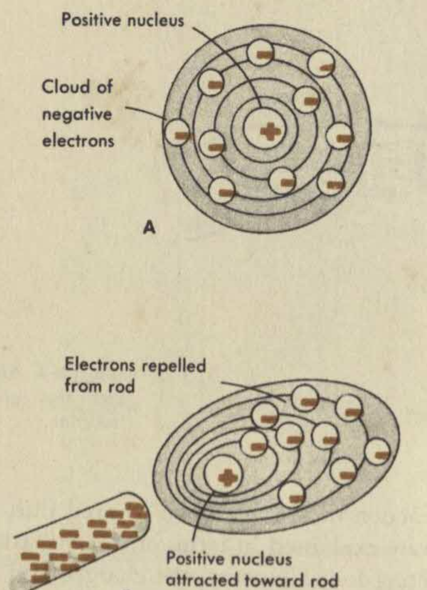


FIG. 12–5 Polarization of an atom in an electric field.

electroscope (Fig. 12-6). It consists of a metal rod, with a right angle bend at the bottom, and a metal ball, which is generally fastened to the top of the rod. A strip of very thin gold foil is draped over the bottom bend of the rod and cemented to it. The rod is supported by a good insulator in the top of a box that surrounds the fragile gold leaves and protects them from damage. When the electroscope is uncharged, the gold leaves hang straight downward from their own weight (Fig. 12-6A). When the electroscope has been charged, however, by having it touch some charged object, the leaves stand apart (Fig. 12-6B) because of the repulsion of the like charges on each leaf.

Figure 12-6C also shows the leaves standing apart, although the electroscope as a whole has no net charge. The positively charged rod near the knob has attracted electrons from the leaves up into the knob, thus giving the leaves a positive charge so that they repel each other. In this case, when the rod is removed, the electrons will return to the leaves and they will again hang down normally, as in Fig. 12-6A.

But if instead of immediately taking the charged rod away, we first touch the knob, the situation will be the same as that shown in Fig. 12-4: the rod will attract more electrons from the finger into the rod. Now if we first remove the touching finger (so that the electrons cannot

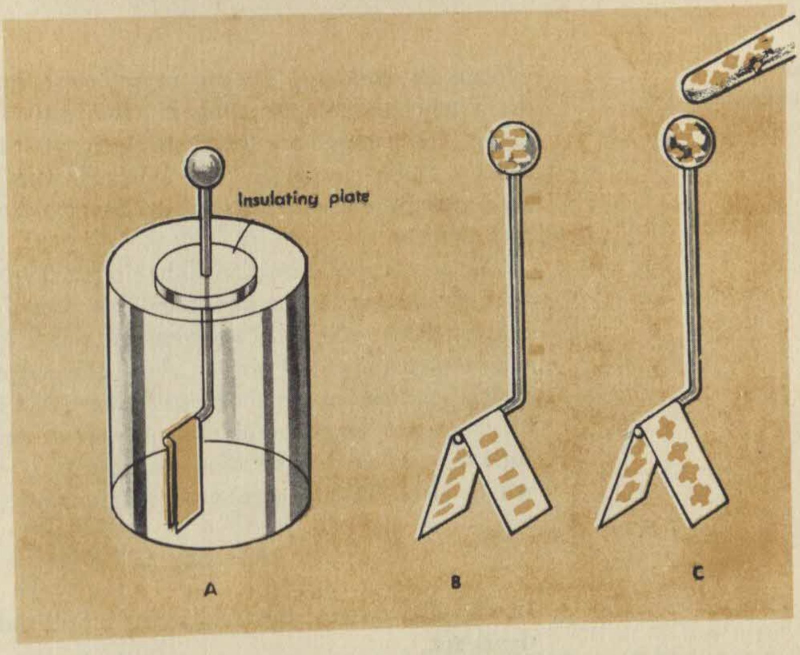


FIG. 12–6 The gold-leaf electroscope.

flow back) and then take the rod away, the electroscope will be left with a charge *opposite* that of the rod. Charging an electroscope by induction in this way is generally more satisfactory than charging it by direct contact.

12-6 The Electric Field

In considering the mechanical interactions between material bodies, we are accustomed to the fact that such interactions require immediate bodily contact. If we want to move an object, we have to touch it with our hand or else have a stick to push it or a rope to pull it. An exception to this need for direct contact is found in the gravitational attraction that all bodies exert on each other. This "action at a distance" bothered Newton when he announced his law of universal gravitational attraction, and he made no attempt to explain *how* gravitational forces were exerted. We say that every mass is surrounded by a *gravitational field* which in some way appears to exert an attractive force on other masses. Einstein has contributed a great deal to our understanding of the gravitational field, but there is nevertheless still much mystery connected with forces which act at a distance.

The forces of electrical repulsion and attraction, acting without any apparent connection between charged bodies, are similar examples of action at a distance and can be attributed to the existence of an *electric field* surrounding charged bodies. Without making any attempt to

explain its workings, we can nevertheless describe an electric field accurately and understandably in terms of the force it exerts on a "test charge." The imaginary test charge to be used is, in the CGS system of units, a *positive* charge of 1 esu. Wherever this test charge is placed in an electric field, it will have exerted on it a force of a certain magnitude and direction.

The strength and direction of the electric field at any point are defined to be the same as the magnitude and direction of the force exerted on our unit test charge when it is placed at that point. In other words, when we use a test charge of $+1$ esu, the electric field strength E equals the force F which the field exerts on the test charge, and is in the same direction. Of course, we can make this concept more general by imagining that we use a test particle bearing some other charge, Q esu. In this case the force will be just Q times as great as for a unit charge, and we can write

$$F = EQ \quad \text{or} \quad E = \frac{F}{Q}.$$

In the CGS system, the field strength will naturally be expressed in dynes/esu.

It is easy to derive a general expression for the strength of the electric field caused by a charged sphere, or by a charged particle of any shape that is small enough to be considered a point. Let us take a small sphere bearing a charge Q and calculate the field at point A, a distance r from it. We put the $+1$ esu standard test charge at this point (Fig. 12-7) and figure the force F exerted on it. Coulomb's law gives this as

$$F = \frac{Q \times 1}{r^2}.$$

Since we always use $+1$ esu as the test charge, we can omit the 1 from the numerator; and since this force per unit charge is the definition of E, the strength of the field, we have simply

$$E = \frac{Q}{r^2}.$$

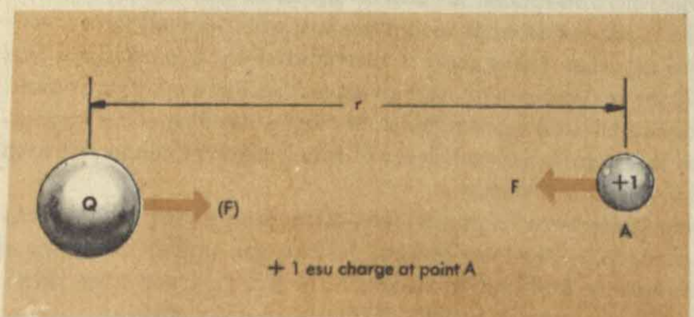


FIG. 12-7 Determination of the electric field at point A, at a distance r from a charged particle.

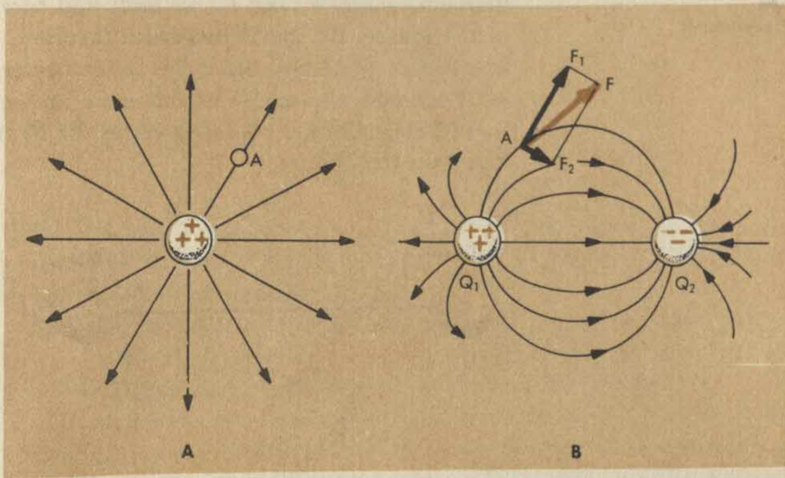


FIG. 12–8 The electric lines of force that map the electric field around charged bodies.

The field can be mapped by drawing lines that at every point run in the direction of the field. Figure 12-8 shows the electric field mapped out with "lines of force" for two simple cases. In Fig. 12-8A, there is a single positively charged sphere, and no matter where the test charge is placed, the force on it will be radially outward away from the central charge in the field we are mapping. Figure 12-8B shows the field resulting from a positive charge Q_1 and a negative charge Q_2 in the same neighborhood. Our $+1$ esu test charge at A will be repelled by Q_1 and attracted to Q_2. The total force on the test charge will be F, the vector sum of F_1 and F_2. The line of force passing through point A, since it is required to give the direction of the field, must be tangent to the vector F.

There is no limit to the number of lines that can be drawn, but it is worth noticing that near the charged bodies, where the field is stronger, the lines are close together; at a distance, where the field is weak, the lines are far apart. Although the drawings show the field only in the two-dimensional plane of the page, the field exists, of course, in three dimensions. It can be shown that in this three-dimensional picture the number of lines of force cutting through a unit area at right angles to the field is exactly proportional to the strength of the field.

Let us take, as an example, four small charged bodies arranged in a square as shown in Fig. 12-9A. What is the electric field at C in the center of the square? We shall determine the field by placing the customary imaginary $+1$ esu test charge at C. It will be subject to four separate forces, one from each of the corner charges. These forces are shown as vectors on the diagram. Because of the symmetry of the arrangement, F_1 and F_3 are equal and opposite in direction and cancel out. To

determine forces F_2 and F_4, we must find how far C is from the corners of the square. By the Pythagorean theorem, a diagonal of the square is $\sqrt{200}$, or 14.14 cm. Since C is at the center, it is 7.07 cm from 2 and 4. (We could also arrive at this same answer by taking $10 \times \sin 45° = 7.07$ cm.) Coulomb's law gives us for F_2, the force between charge 2 and the test charge,

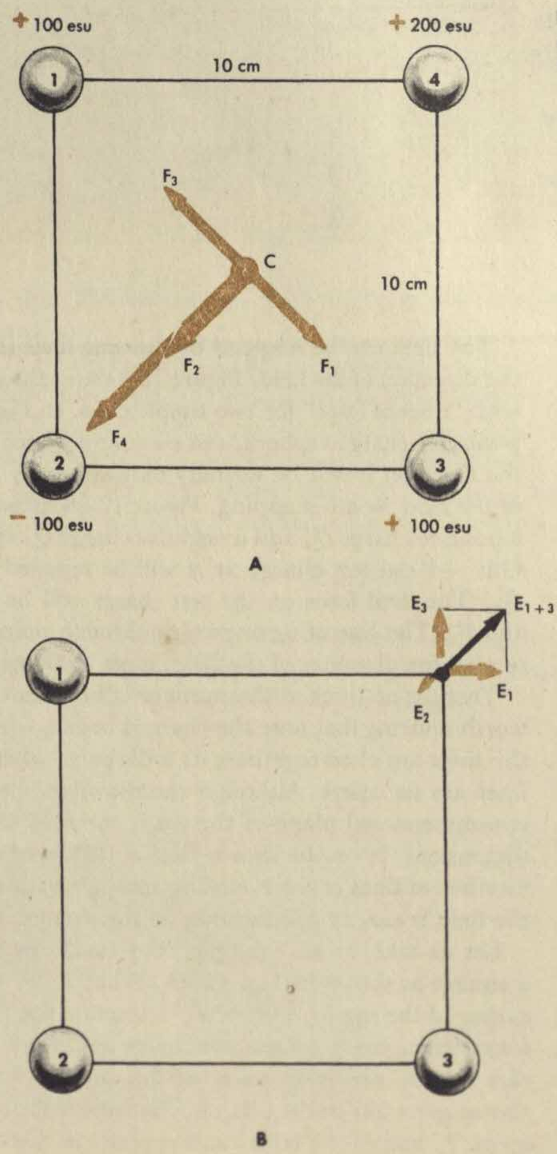


FIG. 12-9 The electric field produced by several charged bodies.

$$F_2 = \frac{Q_1 Q_2}{d^2} = \frac{100 \times 1}{(7.07)^2}$$

$$= \frac{100}{50} = 2 \text{ dynes.}$$

Similarly, we have

$$F_4 = \frac{200}{50} = 4 \text{ dynes.}$$

Since this total force of $2 + 4 = 6$ dynes was calculated on the basis of a charge of $+1$ esu placed at C, this tells us that the field at C is 6 dynes/esu, directed toward charge 2.

Since electric field strength is by definition the electric force on a single esu of charge, it follows that if we put a charge of, say, 7 esu at C, it would experience a force 7 times greater, or 42 dynes.

We can follow up this idea by finding what force will be exerted on charge 4. Charge 4's own field cannot move charge 4 any more than a man can lift himself from the floor by pulling on his own bootstraps. So let us temporarily remove charge 4 (in imagination, at least), find the field there due to the other three charges, then put 4 back. Figure 12-9B shows the other three charges and the components of the field which each one produces. We are able now to omit the customary test charge and figure each field component directly. Components E_1 and E_3 are each $100/10^2 = 1$ dyne/esu. Vectorially, these add to give $E_{1+3} = 1.414$ dynes/esu, at a 45° angle, as shown. Directed toward corner 2, E_2 is $100/200 = 0.5$ dyne/esu, at 45° downward to the left. The sum of all three components of the field (since E_2 and E_{1+3} are along the same line) is $1.414 - 0.5 = 0.914$ dyne/esu. When we replace charge 4 in this field, we find the force on it to be $0.914 \times 200 = 183$ dynes, away from the square along the diagonal line.

12-7 Electric Potential

Imagine a small body bearing an electric charge of Q esu; at a point r cm distant from this charged body, the electric field will have a strength of

$$E = \frac{Q}{r^2}$$

as we saw in the previous section. This same point in the electric field will also have another property called *electric potential. The electric potential at a point is the work needed to bring a $+1$ esu test charge to the point from an infinite distance away.*

We do not need to start from the beginning to calculate an expression to determine this amount of work, because we have already done so on page 98.

Turn back to the section on gravitational potential; there we started

with a force of gravitational attraction, $F_{grav} = GMm/r^2$. The expression for electric attraction (or repulsion) is a similar inverse square law: $F_{elec} = Q_1Q_2/r^2$. If, in the gravitational case, we let m be a 1-gm "test mass," the work needed to remove the test mass from an original distance r from another mass M, to an infinite distance away, would be GM/r. Conversely, the work needed to bring the unit test mass from infinity to r would be merely the negative of this: $-GM/r$. So we could therefore describe the gravitational potential at a distance r from a mass M as $-GM/r$. In an analogous treatment of the electrical case, we let Q_2 be the $+1$ esu test charge and conclude that the electrical potential V at a distance r from a charge Q is given by the simple expression

$$V = \frac{Q}{r}.$$

(The minus sign does not appear here, because charges of the same sign repel, rather than attract, as masses do. If the charge Q were a negative number, the potential V would be negative. The charge Q would attract the positive test charge, which we would have to hold back, rather than push forward as we brought it in.)

It will be a good idea to look carefully at the dimensions of the equation for electric potential we have just derived. The potential V is the work in ergs done on a 1-esu charge, and must therefore have the dimensions of ergs/esu. If we rewrite the equation with dimensions, we have

$$V\left(\frac{\text{ergs}}{\text{esu}}\right) = \frac{Q\ (\text{esu})}{r\ (\text{cm})}.$$

If an equation is correct, the dimensions must be the same on both sides, and this at first glance looks very unlikely in the equation just written. In order to check it, we must first find the dimensions of the esu, which you will remember was defined by Coulomb's law:

$$F\ (\text{dynes}) = \frac{Q\ (\text{esu}) \times Q\ (\text{esu})}{d^2(\text{cm}^2)}.$$

From this we get that $(\text{esu})^2 = \text{dynes} \times \text{cm}^2$, or $\text{esu} = \text{dyne}^{1/2} \times \text{cm}$. If you recall that an erg is a dyne-cm, it is easy to substitute these dimensions and show that the equation is indeed dimensionally correct.

As an example, take a foil-covered pith ball 2 cm in diameter, with a charge of 3 esu, and determine the potential 5 cm from its surface. Here, with the uniformly spherically-distributed charge, its effect will be the same as though all the charge were compressed into a point at the ball's center, 6 cm from where we want to determine the potential. We thus have

$$V = \frac{3\ \text{esu}}{6\ \text{cm}} = 0.5\ \text{erg/esu}.$$

The potential at the ball's surface, only 1 cm from the center, would similarly be found to be $V = 3$ esu/1 cm $= 3$ ergs/esu.

The concept of electric potential is a very important one, and one for which we shall have many uses. Few of these uses will have any apparent connection with the idea of moving a test charge up to a certain distance from a charged body, as we have just done. The procedure of moving up the test charge, however, serves to illustrate and emphasize the important fact that *electric potential is work per unit charge.*

12-8 Practical Electrical Units

So far, we have been working with CGS units (the centimeter, the dyne, the erg), in terms of which we have described electric field intensity and electric potential. But most work with electricity is done in a different system of units, based primarily on the MKS system, and it will be a good idea to switch to these other units before we go further. The "practical" electrical units are based on the coulomb as the unit of charge; the force and work units used are the newton and the joule. *One coulomb equals 3×10^9 electrostatic units of charge.* (This seems to be a strange ratio to pick, but there are good reasons for it. Actually, the number 3×10^9 is one-tenth of the speed of light in centimeters per second; its connection with electric charges comes through the theoretical work of the nineteenth-century British theoretical physicist James C. Maxwell on the propagation of electromagnetic radiation, via still a third system of units with which we need not concern ourselves.)

In practical units, electric field strength is measured by the force in newtons that would be exerted on a hypothetical test charge of $+1$ coulomb, i.e., in units of newtons/coulomb. Electric potential is measured in terms of the familiar *volt. A volt is a joule per coulomb*; as before, the unit of potential is defined as work per unit charge.

Actually, we are seldom concerned with the absolute potential of a point, that is, with the work needed to bring a unit charge up to the point from infinity. More often, we shall deal with the *potential difference* between two points, which is the work required (either positive or negative) to move a coulomb from one point to the other. When we speak of "a 6-volt battery," we mean that the potential difference between the terminals of the battery is 6 volts, i.e., that a coulomb of charge will do 6 joules of work in flowing from one terminal to the other.

12-9 Capacitance

When we originally defined electric potential, it was associated with the presence of an electrically charged body. The surface of a charged conducting ball must be at some definite potential, because it will take a definite amount of work to bring a unit charge up to the ball.

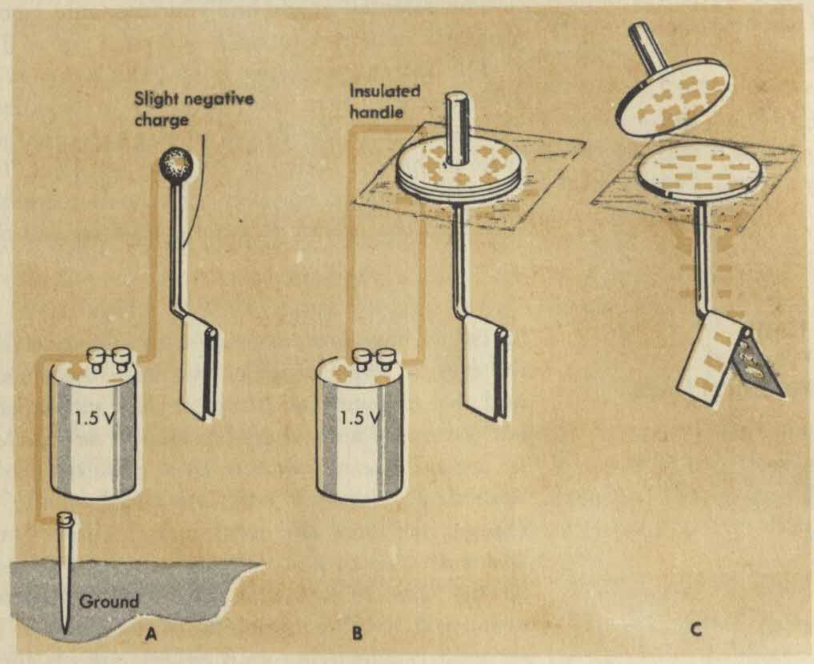


FIG. 12-10 Increasing the charge on an electroscope by using a capacitor.

If the charge on the ball were doubled, it would take twice as much work to bring up a unit charge, and the potential of the ball would accordingly be twice as high. We thus see that the charge on a body is proportional to its potential, and the ratio of charge to potential is called the *capacitance* of the body. The basic unit in which capacitance is measured is the *farad* (f), which is a coulomb per volt. That is, if one coulomb of charge added to a body gives it a potential of one volt, it has a capacitance of one farad:

$$C \text{ (farads)} = \frac{Q \text{ (coulombs)}}{V \text{ (volts)}}.$$

The farad is an enormously large unit, and capacitance is more often measured in microfarads (1 microfarad $= 10^{-6}$ farad, abbreviated μf) or even micromicrofarads (10^{-12} farad, $\mu\mu$f, now often referred to as a *picofarad*, pf).

The capacitance of an isolated body such as an electroscope is very small. If an ordinary dry cell (about 1.5 volts) is connected between the electroscope knob and the ground (Fig. 12-10A), thereby moving electrons from the ground to the knob until the electroscope has a potential of 1.5 volts with respect to the ground, there will be no perceptible move-

ment of the leaves. The Q transferred to the electroscope is so small that it, will cause no perceptible repulsion of the leaves.

We can modify the electroscope by replacing the knob with a large flat sheet of metal. On this we put a thin sheet of paper, cellophane, or other insulating material and complete the sandwich with another flat metal sheet, as shown in Fig. 12-10B. The battery will pull electrons from the top plate through the battery and into the bottom electroscope plate until the potential difference between the two plates reaches 1.5 volts. The amount of charge transferred will be enormously greater than in Fig. 12-10A, however. The attraction of the positively charged top plate will help bring in and hold the negative charge on the bottom plate. This attraction of the top plate will also prevent much negative charge from accumulating on the gold leaves, and the leaves will hang nearly straight down. The battery can now be disconnected. (The charging of the two plates will take only a fraction of a second.) When the top plate is removed, the mutual repulsion of the excess electrons on the electroscope will drive some of the charge down into the leaves, which will now stand apart (Fig. 12-10C).

A device such as we have put on the electroscope—two sheets of metal or other conductor, with a sheet of insulating material between them—is called a *condenser* or *capacitor*. The capacitance of the condenser—that is, the amount of electrical charge on its plates divided by the potential difference between its plates—depends on several factors. One such factor is obviously its area; if the plates are made twice as large, the charge held on them will be twice as great. The thickness of the insulating layer between the plates is also important. The closer the plates are to one another, the greater is the amount of charge that is held, because of the increase in the strength of the electric field between the plates as they are brought closer together.

FIG. 12–11 Charged parallel plates with vacuum (or air) between them.

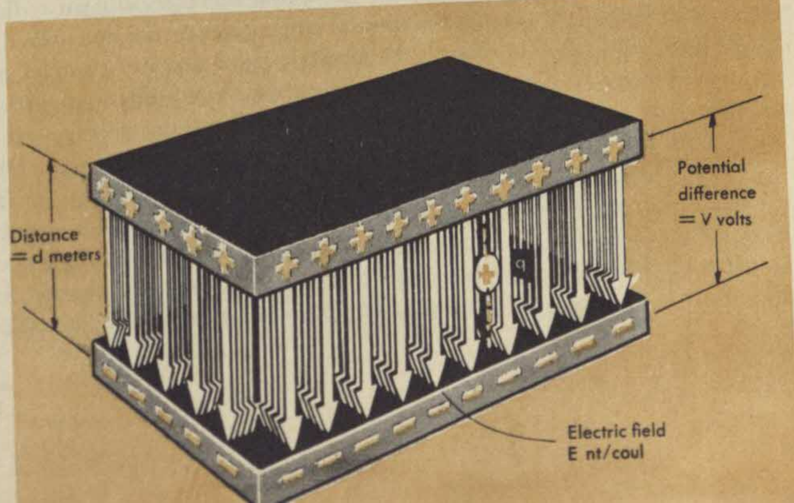


Figure 12-11 shows a pair of parallel charged plates separated by d meters, and with a potential difference of V volts between them. In the empty space between them will be an electric field with a strength we can call E nt/coul. This field will be uniform and perpendicular to the plates except at the very edge. In order to find the strength of the field, let us take a small charge of $+q$ coulombs and move it from the negative plate to the positive plate. The work needed to do this is $F \times d$, or Eqd joules. The work can also be found in another way: since V is by definition the work in joules needed to move 1 coulomb from one plate to the other, then to move q coulombs will require Vq joules. Equating the two expressions for the same work, we get

$$Eqd = Vq$$

or

$$E \text{ nt/coul} = \frac{V}{d} \text{ volts/meter.}$$

Thus we see that expressing electric field strength in volts/meter is exactly numerically equivalent to expressing it according to its definition as newtons/coulomb. The closer together the plates are, the smaller d is, and with the same potential difference V, the stronger will be the field and the more charge will be held on the plates. Further analysis shows that Q is exactly inversely proportional to d.

If, for example, we have a parallel-plate capacitor made of two metal sheets 10 cm $\times$ 10 cm, 2 mm apart, and connect the plates to the terminals of a 45-volt battery, we have for V/d: $45/2 \times 10^{-3} = 22{,}500$ volts/meter. The electric field between the plates in thus 22,500 nt/coul.

The material between the plates also has a strong influence on the capacitance of a condenser. Figure 12-12 shows a dielectric slab between the plates. (Insulators or nonconductors are known as *dielectrics*.) We have already mentioned the distortion or *polarization* of atoms in an electric field, and in Fig. 12-12 the atoms of the dielectric between the plates are represented schematically in a distorted form. The positive signs represent the nuclei thrust downward by the field between the plates; the negative signs above the nuclei represent the centers of the clouds of electrons which are thrust upward. In the central material of the slab, these atomic distortions average out to no net effect. At the top surface, however, the upward-pulled electrons result in a layer of excess negative charge. Similarly, on the bottom face of the slab, the downward-pulled positive nuclej produce a surface layer of positive charge. These layers of induced charge help maintain the charges on the metal plates and result in a higher capacitance for the cndenser. For our purposes, we can define the *dielectric constant* of an insulating material as

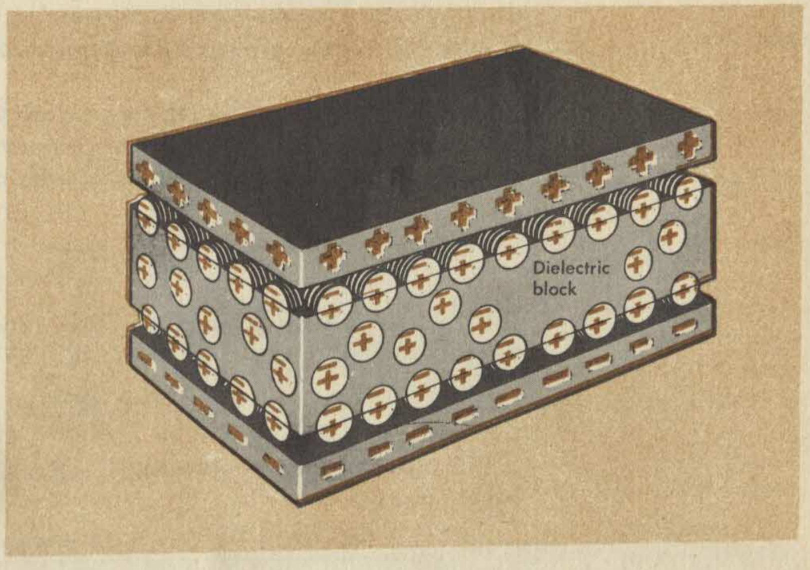


FIG. 12-12 Charged parallel plates separated by a dielectric material.

the factor by which it multiplies the capacitance of a condenser, as compared with the same condenser with a vacuum between the plates. Here are dielectric constants K for a few common dielectrics:

Hard rubber	2.5
Glass	6 to 10
Oil	2 to 2.3
Water (pure)	80
Mica	5.5 to 6.5
Air	1.001

What causes the difference in the dielectric constants for different materials? Why does a sheet of glass between the plates of a capacitor increase its capacitance by a factor of, say, 6, while the same capacitor dipped into oil would have its capacitance only doubled? The answer is, of course, that some atoms or molecules are more readily polarizable than others. In glass, the layers of + and − charge on each side are stronger and better separated than the corresponding layers formed in oil. Water, with its enormous dielectric constant of 80, introduces a new factor, because the water molecule is already polarized, even if it is not in an electric field. One end of a water molecule is positive, and the other negative; in an electric field, the molecule merely rotates (which it can

readily do in its liquid state) in response to the electric forces on it, and thereby readily produces strong layers of induced charge on its surfaces.

Putting all this discussion together, we see that the capacitance of a parallel-plate capacitor is proportional to its area A and to the dielectric constant K, and inversely proportional to the distance d separating the plates:

$$C = k \times \frac{KA}{d}.$$

If C is to be in farads, A in square meters, and d in meters, the proportionality constant k can be shown to be 8.85×10^{-12}, which gives us

$$C = \frac{8.85 \times 10^{-12} KA}{d}.$$

The capacitor used as an example earlier in this section thus has a capacitance of

$$C = \frac{8.85 \times 10^{-12} \times 1 \times 0.10 \times 0.10^{*}}{2 \times 10^{-3}} = 4.42 \times 10^{-11} \text{ f}$$
$$= 44.2 \text{ pf.}$$

The charge on each of the plates of this capacitor at a potential difference of 45 volts would be

$$Q = CV = 4.42 \times 10^{-11} \times 45 = 1.99 \times 10^{-9} \text{ coul.}$$

If a sheet of hard rubber 2 mm thick were slipped between the capacitor plates, with the battery still connected, V would of course remain the same, but the capacitance and the charge on the plates would each become greater by a factor of 2.5. Suppose, though, that the battery is first disconnected, and then the rubber sheet slipped into position. Since there is no conductor through which electrons can move, the charge on the plates must remain the same. The capacitance is again increased by a factor of 2.5; and since $V = Q/C$, the potential difference between the plates drops to 18 volts.

* If the dimensions were all accurate to one part in a thousand and we wanted an answer to equal accuracy, it would have been necessary to use 1.001 for the dielectric constant of the air between the plates. Since only a very few special capacitors would warrant calculations this precise, the difference between vacuum (1.000) and air (1.001) has been neglected here.

Questions

(12-1) 1. Two small charged balls are placed 5 cm apart in air. One ball has a charge of $+6$ esu, the other -20 esu. What is the force between them?

2. Two small charged balls, 8 cm apart in air, have charges of $+12$ and $+4$ esu. What is the force between them?

3. Two small equally charged balls repel each other with a force of 12 dynes when they are 4 cm apart. How much charge is on each of the balls?

4. Two protons repel each other with a force 4×10^{-6} dyne when they are 2.4×10^{-7} cm apart. What is the magnitude of the charge on each proton?

5. An electron has a charge of 4.8×10^{-10} esu; the charge of an alpha particle is twice as great, and positive. How far apart are an electron and an alpha particle when there is an attraction of 6×10^{-6} dyne between them?

6. One small ball has a charge of $+6$ esu; another has a charge of -12 esu. How far apart must they be to attract each other with a force of 4.5 dynes?

(12-3) 7. Two charged balls are identical and covered with a metal coating. They attract each other with a force of 27 dynes when they are 4 cm apart. The balls are touched together (when such conducting balls are touched together, their total charge divides equally between them) and again placed 4 cm apart; they now repel each other with a force of 9 dynes. (a) What is the explanation for this switch from attraction to repulsion? (b) What is the charge on each ball after touching? (c) What were the charges before touching?

8. Two identical metal-covered balls attract each other with a force of 8 dynes when they are 3 cm apart. The balls are touched together (see Question 7) and again placed 3 cm apart; they now repel each other with a force of 1 dyne. (a) What is the explanation for this switch from attraction to repulsion? (b) What is the charge on each ball after touching? (c) What were the charges before touching?

(12-4) 9. In previous questions, we have tacitly assumed that the conductive balls were small enough in comparison to their distance apart that any migration of charge over their surfaces would not affect our answers. Consider, however, 2 such balls 0.5 cm in diameter, with a charge of -2 esu on each. Place the balls 2 cm apart, measured center-to-center. Taking the inevitable migration of charge into account, would the actual repulsion between them be greater or less than the 1 dyne we might have expected?

10. Repeat Question 9 for the case in which the charges on the balls are $+2$ and -2 esu. Will their attraction be greater or less than 1 dyne?

(12-5) 11. Given a neutral electroscope, a bit of fur, and a rubber rod, describe two ways to give the electroscope a positive charge, and two ways to give it a negative charge.

12. Given a neutral electroscope, a piece of silk, and a glass rod, describe two ways to give the electroscope a positive charge, and two ways to give it a negative charge.

13. An electroscope has been charged by touching the knob with a piece of glass that has been rubbed with silk. As a charged object is brought near the

electroscope, the leaves come closer together until they hang straight down. When the object is brought still closer (but without ever touching the electroscope), the leaves stand apart again. (a) Explain what has happened during this experiment. (b) Is the charge on the object positive or negative?

14. A hard rubber rod that has been rubbed with fur is brought near the knob of an already charged electroscope. As it approaches, the leaves come closer together until they hang straight down. When the rod is brought still closer (but without touching the knob) the leaves stand apart again. (a) Explain what has happened during the experiment. (b) Was the charge on the electroscope positive or negative?

(12-6) **15.** What is the intensity and direction of the electric field at a point 10 cm north of a small body bearing a charge of $+120$ esu?

16. Point A is 5 cm west of a particle having a charge of -75 esu. What is the magnitude and direction of the electric field at A?

17. An alpha particle has a charge of $+9.6 \times 10^{-10}$ esu. (a) What is the electric field 3×10^{-7} cm from an alpha particle? (b) What is the force on an electron at this distance from an alpha particle? (An electron has a negative charge just half as large as the charge on a alpha particle.)

18. (a) What is the electric field 3×10^{-7} cm from an electron? (See Question 17.) (b) What is the force on an alpha particle at this distance from an electron?

19. A body bearing a charge of $+64$ esu is 12 cm from another body that has a charge of $+256$ esu. At what point (or points) is the electric field zero?

20. A particle with a charge of -7 esu is 6 cm from another particle with a charge of -63 esu. At what point (or points) is the electric field zero?

21. What is the magnitude of the force on an electron that is 3×10^{-7} cm from one alpha particle, and 4×10^{-7} cm from another, the alpha particles being 5×10^{-7} cm apart? (See Question 17.)

22. Point A is 10 cm from point B; C is 8 cm from B, and 6 cm from A. Charges of 288 esu are at each of the 3 points. What is the magnitude of the force on the charge at point C?

(12-7) **23.** What is the potential at the point described in Question 15?

24. What is the potential at point A in Question 16?

25. How much work would be required to move the electron from the position described in Question 17 to a great distance away?

26. How much work would be required to remove an alpha particle from point A in Question 16?

27. (a) In the situation of Question 19, is there a point at which the potential is zero? (b) What if we make the 64-esu charge negative? If there is now a zero potential point (or points) on the straight line through the charges, where is it located?

28. (a) Is there a point of zero potential near the charges of Question 20? (b) If the 7-esu charge is made positive instead of negative, is there now a zero potential point (or points) on the straight line through the charges? If so, where?

(12-8) **29.** A charge of $+10^{-7}$ coulomb ($= 0.1\ \mu$ coul, or 0.1 microcoulomb) is 15 cm from a charge of $-0.2\ \mu$ coulomb. (a) What is the potential (in volts) at A, 5 cm from the positive charge, and on the line connecting the two charges? (b) What is the potential at B, 4 cm from the negative charge and on the same line? (c) What is the potential *difference* (in volts) between A and B? (d) How much work (in joules) would be required to move a charge of $+1\ \mu$ coulomb from B to A?

30. A charge of -3×10^{-8} coulomb is 20 cm from a $+9 \times 10^{-9}$ coulomb charge. Point A is on the line joining the charges, and is 6 cm from the negative charge; point B is on the same line, 3 cm from the positive charge. (a) What is the potential (in volts) at A? (b) What is the potential at B? (c) What is the potential *difference* (in volts) between A and B? (d) How much work would be required to move a charge of -2×10^{-8} coulomb from A to B?

(12-9) **31.** A capacitor formed of a pair of parallel metal plates with air between them has a capacitance C. What will be its capacitance if the plates are brought to one-fourth as far apart as they were, and the space between them filled with mica?

32. A capacitor consists of a pair of parallel metal plates with oil (dielec. const. = 2) between them and has a capacitance C. What will its capacitance be if we drain the oil from between them, and make their separation three times what it originally was?

33. A 10^{-11} f capacitor is made of two flat parallel metal plates 1 cm apart. By means of a battery a potential difference of 100 volts is maintained between them. (a) What is the electric field between the plates? (b) What is the electric field between the plates when they are brought to 0.5 cm apart?

34. A 10^{-11} f capacitor is made of two flat parallel metal plates 1 cm apart. By connecting them to battery terminals, the plates are charged to a potential difference of 100 volts. The battery is then *disconnected*, so that the charges on the plates must remain unchanged. (a) What is the electric field between the plates? (b) What is the electric field between the plates when they are brought to 0.5 cm apart?

35. In Question 34, what is the potential difference between the plates when they are brought to 0.5 cm apart?

36. In Question 33, what is the charge on each plate (a) when they are 1 cm apart? (b) when they are brought to 0.5 cm apart?

37. A parallel-plate capacitor remains connected to a battery. Is the charge on the plates increased or decreased when a sheet of glass is substituted for the air separating them?

38. Battery terminals are briefly connected to the plates of a capacitor, and are then disconnected. Does the electric field between them become stronger or weaker when the air separating the plates is replaced by oil?

39. After the battery of Question 34 is disconnected, the 1-cm space between the plates is filled with oil whose dielectric constant is 2.2. What, then, is the potential difference between the plates?

40. What is the charge on the plates of the capacitor of Question 33 when the 1-cm space between them has been filled with oil of dielectric constant 2.0?

41. What is the capacitance (in μf) of a capacitor made of two metal sheets 10 cm $\times$ 20 cm, separated by 2 mm of glass whose dielectric constant is 7.5?

42. A capacitor consists of two sheets of aluminum foil 3 cm $\times$ 10 m, separated by a strip of waxed paper (diel. const. 2.5) 0.1 mm thick. What is the capacitance, in μf?

chapter / thirteen

Electric Currents

13-1 Electric Cells

If we connect a conductor charged with, let us say, positive electricity, to the ground or to a negatively charged conductor by means of a metallic wire (Fig. 13-1A), the conductor will be discharged. For a split second, a current of electrons will flow through the wire. The duration of such an electric discharge is, however, too short for convenient study, and it is desirable to have an arrangement that provides us with a steady current. This became possible after the discovery that steady electric current can be produced by the *electric cell*, invented by the Italian physicist Alessandro Volta (1745–1827). The simplest sort of voltaic cell consists of two plates or rods (called *electrodes*) made of two different materials, such as carbon and zinc, placed in some conducting solution (called an *electrolyte*), such as sulfuric acid (Fig. 13-1B).

To study in detail what goes on in an electric cell or battery would take us far afield into chemistry. However, a few elementary ideas will give a fair overall picture of the process by which the cell converts chemical energy into electrical energy. The first of these ideas concerns *ioniza-*

tion. An *ion* is an atom or a group of atoms which has either an excess or a deficiency of electrons and thus carries a net negative or positive charge. Many substances—salts, acids, and bases—ionize when they are dissolved in water. This means that the molecules of the substance separate into pieces which do not share the electrons evenly, so that instead of an electrically neutral molecule we have two or more pieces, each with an electric charge.

With such electrically charged ions floating freely about in it the solution is able to conduct an electric current and is an electrolyte. The electric cell mentioned above was made by putting carbon and zinc into a sulfuric acid solution. The chemical formula for sulfuric acid is H_2SO_4, which means it is composed of 2 hydrogen atoms, 1 sulfur atom, and 4 oxygen atoms. When sulfuric acid is dissolved in water, the hydrogens split off from the rest of the molecule, each hydrogen leaving an electron behind, and the molecule thus becomes 3 ions. Each of the hydrogens, having left an electron behind, has an excess positive charge; the SO_4 remaining (called the "sulfate" group), having fallen heir to the extra electrons, has 2 excess negative charges. We can denote this ionization in the following way:

$$H_2SO_4 \longrightarrow H^+ + H^+ + SO_4^{--}.$$

The metal zinc is not soluble in water. However, when it is placed in an acid solution, some of the zinc atoms become zinc ions, which are

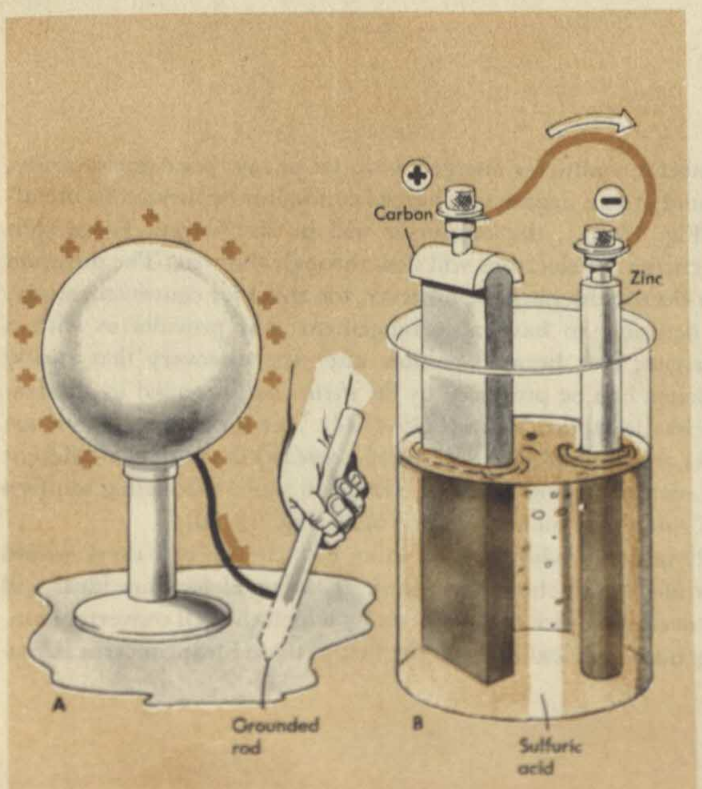


FIG. 13–1 (A) An instantaneous electric current from a charged conductor; (B) a continuous electric current from a chemical cell.

soluble. The zinc atom ionizes by leaving electrons behind, as do all metals. Each zinc atom which goes from the metal plate into the acid solution leaves 2 electrons behind on the plate and thus becomes a doubly charged positive ion:

$$Zn \longrightarrow Zn^{++} + e^- + e^-.$$

Let us consider how the cell of Fig. 13-lB operates to produce a steady flow of electrons through the wire connecting the zinc to the carbon. A Zn^{++} goes into solution from the zinc rod, thereby leaving the rod with an excess of 2 electrons. The electrolyte must remain electrically neutral, so for each Zn^{++} that goes into solution, two H^+'s leave the solution by picking up electrons from the carbon rod and becoming H atoms. The neutral H atoms are not soluble and rise to the top of the electrolyte in bubbles of hydrogen gas. The robbing of electrons from the carbon rod has left it with a deficiency of 2 electrons, and the excess electrons left behind by the zinc flow along the wire to the carbon to make up the deficiency. This sequence of events happens billions or trillions of times a second, with the result that there is a steady electron current flowing through the wire.

Any pair of dissimilar conducting materials can be put into any electrolyte and will become a cell which will convert chemical energy into electrical energy in a manner similar to the cell described above. These cells will vary widely, however, in the amount of chemical energy released per coulomb of charge. You will recognize this last sentence as another way of saying "these different cells will have different voltages," since "voltage" means the energy in joules per coulomb of charge.

13-2 Electrical Resistance

In dealing with the electric cell, we have spoken of the stream of electrons flowing along the wire connecting the negative electrode of the cell (called the *cathode*) to the positive electrode (called the *anode*). This electron flow, or current, could be described as so many electrons per second. However, this would require the use of enormously large numbers, and scientists have found it convenient to invent a special unit to describe the flow of electric charge. This unit, named for a French physicist and early electrical experimenter, André Ampère (1775–1836), is the *ampere* (amp). *One ampere is a current of one coulomb per second.*

Just as water flows faster through a pipe when the difference in pressure between the ends of the pipe is larger, electric current depends on the difference in electric potential between the ends of the wire. The law discovered by a German physicist, G. S. Ohm (1787–1854), states that for many substances, particularly metals, ***the electric current in a conductor is directly proportional to the potential difference between its ends.*** In other words, every conductor presents some resistance to the

passage of an electric current, and to increase a current through it, proportionally more energy per coulomb of charge is required. On the basis of Ohm's law, a unit of electrical resistance has been defined and named the *ohm*. The ohm can be defined as follows: *the resistance of a wire or other conductor is 1 ohm if a potential difference of 1 volt between its ends will cause a current of 1 ampere to flow through it.*

In *metal* conductors, Ohm's law is obeyed, and the electric current I is proportional to the potential difference V and inversely proportional to the resistance of the conductor R. Because of the way the ohm was defined, this relationship can be put into the very brief and convenient form known as Ohm's law:

$$I \quad (\text{amperes}) = \frac{V \quad (\text{volts})}{R \quad (\text{ohms})}$$

or

$$V \quad (\text{volts}) = I \quad (\text{amperes}) \times R \quad (\text{ohms})$$

or

$$R \quad (\text{ohms}) = \frac{V \quad (\text{volts})}{I \quad (\text{amperes})}.$$

Ohm's law does not hold for the current through the electrolytes of cells, for electrical discharge through gases, for vacuum tubes, or for devices which use poorly conducting materials called semiconductors. However, in most electric circuits, the current flows through metallic conductors, and we shall make extensive use of Ohm's law when we take up electrical circuits.

Electrical conductors come in all shapes and sizes—the most common, of course, being a wire, which is merely a long cylinder. In order to describe and compare the abilities of various materials to conduct electric currents, handbooks tabulate their resistivities. *The resistivity of a material is the resistance of a conductor 1 cm long and of 1 cm² cross-sectional area.* (See Table 13-1.)

Intuition tells us that a long conductor should have more resistance than a short one, and that a fat, thick conductor will have less resistance than a thin, fine one. Experimental evidence shows this guess to be a good one: resistance is directly proportional to the length of the conductor and inversely proportional to its cross-sectional area:

$$\text{Resistance} = \frac{\text{resistivity} \times \text{length}}{\text{cross-sectional area}}.$$

It is a perfectly straightforward matter to calculate the resistance of a conductor of any size—a copper wire, for example, 150 m long and 0.086 cm in diameter. Its cross-sectional area is $\pi r^2 = \pi \times (4.3 \times 10^{-2})^2 = 5.80 \times 10^{-3}$ cm²; its length is 1.50×10^4 cm:

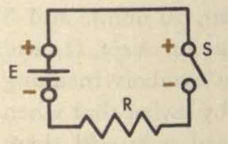
$$R = \frac{1.77 \times 10^{-6} \times 1.50 \times 10^4}{5.80 \times 10^{-3}} = 4.58 \text{ ohms}.$$

TABLE 13-1 ELECTRICAL RESISTIVITIES (AT ROOM TEMPERATURE)

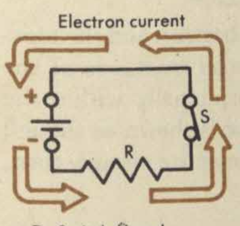
Material	Resistivity (ohm-cm)
Silver	1.63×10^{-6}
Copper	1.77×10^{-6}
Gold	2.44×10^{-6}
Tin	11.5×10^{-6}
Lead	22.0×10^{-6}
Pure water	5×10^{5}
Wood (maple)	3×10^{10}
Glass	2×10^{13}
Amber	5×10^{16}

13-3 Electrical Circuits

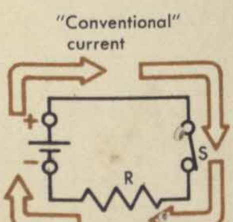
FIG. 13-2 A simple electric circuit.



A Switch open



B Switch closed; flow of electrons



C Switch closed; flow of hypothetical positive charges

A simple electrical circuit has already been shown in Fig. 13-1, and the flow of charge through the circuit has been discussed qualitatively. A circuit must include some source of potential, or energy (in order to make the charge move), and a path through which the charge can travel. Figure 13-2 shows another simple electrical circuit: an electric cell to serve as the energy source, a resistance of some sort (symbolized by the zigzag line—it could be a heater or a light bulb), and a switch. In Fig. 13-2A the switch is "open," and there is no complete path, or circuit, around which the electrons can flow. The cell can only pull electrons from the upper part of the circuit and push them into the lower part until the open contacts of the switch have the same potential difference as the terminals of the cell. This will take only a small fraction of a second, and then there will be no further motion of electrons.

The switch has been closed in Fig. 13-2B, and electrons can now flow around a complete circuit. The direction of electron flow will, of course, be from the negative terminal of the cell, where there is an excess of electrons, counterclockwise around the circuit to the positive cell terminal where there is a deficiency of electrons. Unfortunately for the physics student and his instructor, the "conventional" electric current flows in the opposite direction (Fig. 13-2C). This convention of considering the flow to be from the + terminal through the circuit and back to the − terminal was adopted more than a century before electrons were known to exist and is firmly established in the literature of electricity. Actually (on the principle that two negatives are equivalent to a positive) it makes no difference if we consider hypothetical positive charges moving in one direction, or real negative charges moving in the opposite direction. The conventional direction has the advantage, also, that it is more natural to think of a flow from high potential to low than it is to picture the reverse.

The letter E by the cell stands for *electromotive force* (emf), which is not a force at all but the electric potential produced by the cell. That

is, it is the total energy given each unit of charge as it passes through the cell, as a result of chemical energy being converted into electrical energy. It might well have been labeled V, but since the voltage of the terminals of the cells may be different from its emf, it is well to avoid confusion and reserve E for this use. If R is the total resistance of the circuit and E is the total emf, then the current in the circuit, by Ohm's law, will be $I = E/R$.

13-4 Series Circuits

Figure 13-3 shows a situation a bit more complex. Instead of a single cell, we have a *battery* of five cells connected *in series*. If cells, or resistances, or any other elements in a circuit are in series, it means they are connected together without side branches in such a way that the entire current must flow through each one. The drawing shows not only the cells, but the resistances, all connected in series. If each cell has an emf of 2 volts, it means that 2 joules of potential energy will be given to each coulomb of charge that flows through the cell. Hence a coulomb of charge, which must flow through each of the five cells in series, will be given 2 joules by each cell, for a total of 10. Thus the total emf of the battery is 10 joules/coulomb, or 10 volts.

The current passes through resistances of 2 ohms, 10 ohms, and 8 ohms, in series, for a total of 20 ohms. (The Greek capital *omega*, Ω, and sometimes the small *omega*, ω, are often used as symbols meaning "ohms.") We may summarize the argument above by saying that when emf's or resistances are connected *in series*, we need only add them together:

$$R \text{ total} = \sum R \qquad E \text{ total} = \sum E.$$

The letters *a* and *b* represent the terminals of the battery, and the resistance R_3, lying between the terminals, is the *internal* resistance of the battery. This 2-ohm resistance actually occurs principally within the electrolyte of each of the cells, but for convenience is shown as though it were all gathered together in one place. For the entire circuit, then,

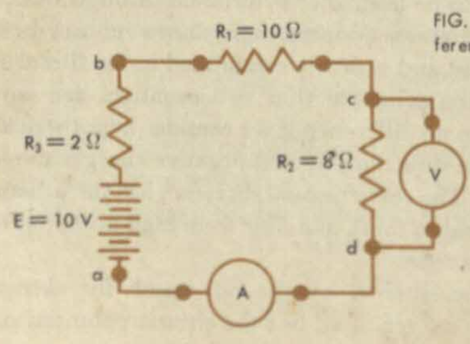


FIG. 13-3 Measurement of potential difference and current in an electric circuit.

we have a total emf of 10 volts and a total resistance of 20 ohms, and hence a current of 10/20, or 0.5 amp.

On the drawing is shown an ammeter A and a voltmeter V. We shall not be able to discuss how they operate until later, but even without a knowledge of their operation some things can be said about their use. The voltmeter is connected to the circuit to measure the difference in potential between points c and d and is not concerned with how much current is flowing. Voltmeters are made with a *very high* resistance, so that only an insignificant trickle of current runs from c to d through the meter, and the original unmetered circuit is virtually unchanged. Forgetting the rest of the circuit for a moment, we see that the voltmeter is connected to read the potential difference between the ends of a resistance through which a current of 0.5 amp is flowing. By Ohm's law we know that this potential difference must be

$$\begin{aligned} V &= IR \\ &= 0.5 \times 8 = 4.0 \text{ volts.} \end{aligned}$$

Because of the way it is figured, we can refer to the potential difference between the ends of a resistance through which a current flows as the "IR drop across the resistance."

Similarly, the IR drop across R_1 is $0.5 \times 10 = 5.0$ volts; across R_3 it is $0.5 \times 2 = 1.0$ volt. The sum of the three IR drops is $4.0 + 5.0 + 1.0 = 10.0$ volts, which is, as it must be, the emf of the battery. The energy given the circulating charge is exactly equal to the energy it expends in heating the resistances through which it flows.

If we were to connect the voltmeter to a and b, the battery terminals, it would not read the emf of the battery, but something less, because some of the energy given to the charge is used up within the battery itself. Work must be done in moving the charge through the internal resistance of the cell. This work, like any work done to overcome resistance, is converted into heat and lost. As we have seen, this internal IR drop inside the battery is $0.5 \times 2 = 1.0$ volt, so we would expect the terminal voltage to be $10.0 - 1.0 = 9.0$ volts.

In a series circuit, it makes no difference where the ammeter is placed, because the current is the same in every part of the circuit. It is connected in a very different way from the voltmeter, however. It must itself be placed in series in the circuit, so all the current will pass through it. The ammeter must therefore have a *very low* resistance in order for its effect on the circuit to be negligible.

13-5 Parallel Circuits

We have already seen an example of a current which divides into two branches; in Fig. 13-3, the 8-ohm resistance and the voltmeter provide two such branches for the main current to flow through, and are

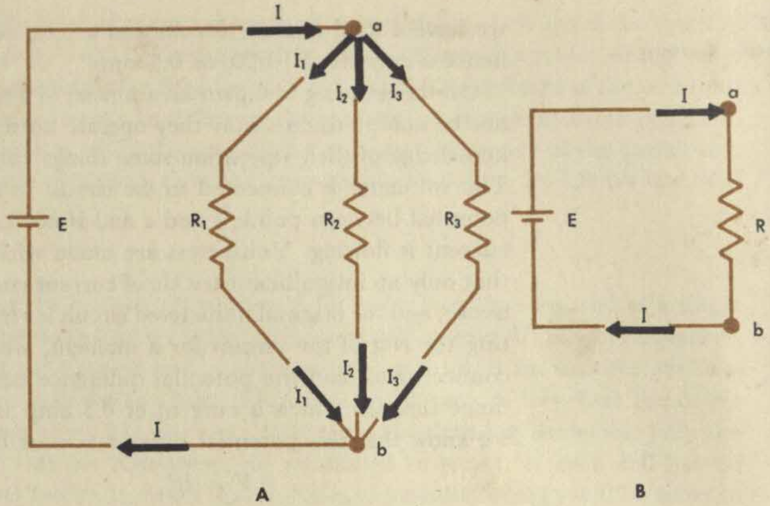


FIG. 13-4 Resistances connected in parallel.

said to be connected *in parallel.* Figure 13-4A shows a circuit in which three resistances are connected in parallel; when the current I reaches point a, it divides into three separate streams, I_1, I_2, and I_3, which recombine at b and return to the negative terminal of the battery. (It is a well-established convention that in drawing a cell or battery the long thin cross-bar is $+$ and the short thick cross-bar is $-$.) We want to compute the value of the one single resistance R (in Fig. 13-4B) that can be substituted for R_1, R_2, and R_3 and still have the same current I and the same potential difference between a and b.

In the arrangement shown in the drawing, the potential difference between a and b (or, as it is often stated, the potential across a and b) we shall call V. We can now write three simple equations for the separate branch currents:

$$I_1 = \frac{V}{R_1} \qquad I_2 = \frac{V}{R_2} \qquad I_3 = \frac{V}{R_3}.$$

Turning to Fig. 13-4B, we can similarly write for the main current

$$I = \frac{V}{R}.$$

It is apparent that the sum of the three branch currents must equal the main current:

$$I = I_1 + I_2 + I_3$$

or

$$\frac{V}{R} = \frac{V}{R_1} + \frac{V}{R_2} + \frac{V}{R_3}.$$

Dividing this entire equation by V, we get

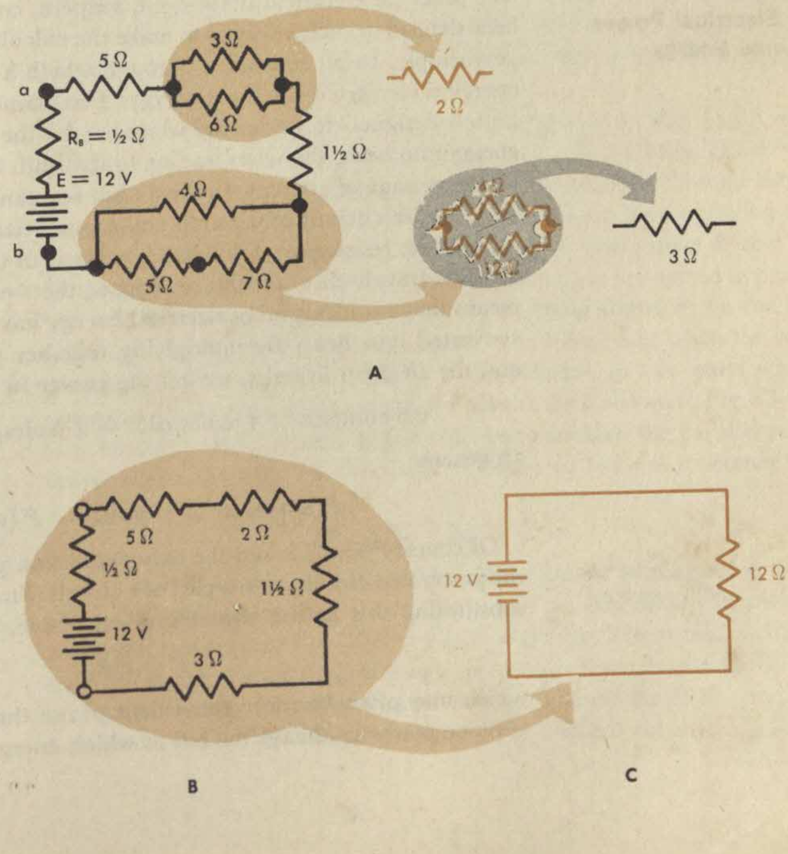
$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

which is the relationship we need to be able to calculate the equivalent resistance R from the values of the resistances connected in parallel. An inspection of the derivation will show that this relationship is applicable not only to three but to any number of resistances in parallel.

Figure 13-5 shows the steps followed in simplifying a complex circuit into an equivalent elementary circuit containing one emf and one resistance. In general, it is a good idea to get rid of any parallel branches, by replacing them with the single resistances to which they are equivalent. In Fig. 13-5A, we see there are two such places in the circuit. In the top one, 3 ohms and 6 ohms are in parallel. To find the single resistance with which they can be replaced, we write

$$\frac{1}{R} = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}; \qquad R = 2 \text{ ohms.}$$

FIG. 13–5 The simplification of a circuit containing parallel branches.



In the bottom circle, the bottom branch consists of 5 ohms and 7 ohms *in series*; these can be immediately combined by simple addition, so that what we have is 4 ohms and 12 ohms in parallel. For this equivalent resistance,

$$\frac{1}{R} = \frac{1}{4} + \frac{1}{12} = \frac{1}{3}; \qquad R = 3 \text{ ohms.}$$

Now, in Fig. 13-5B, each of the two parallel-connected parts of the circuit has been replaced by its single equivalent resistance. From here it is easy to merely add together all these series-connected resistances, and arrive finally at the simplest possible circuit of Fig. 13-5C.

(It should be noted that it is possible to have more complex circuits which may contain elements not unambiguously connected either in series or in parallel with other elements. The straightforward procedure used for Fig. 13-5 may not work in such cases, and a more advanced procedure must be used. We shall not deal with such circuits in this book.)

13-6 Electrical Power and Energy

The practical electrical units—volt, ampere, coulomb, and ohm—have been defined in such a way as to make the calculation of work and power very simple. In all resistances through which a current flows, electrical energy is converted into heat energy. The filament of an ordinary light-bulb is designed to be heated white-hot by the conversion of electrical energy into heat; the wires leading to the bulb also are slightly warmed by the passage of current. In the 8-ohm resistance in Fig. 13-3, we have calculated a current of 0.5 amp and a potential difference between the ends of the resistance of 4 volts. A current of 0.5 amp means that 0.5 coul/sec travels through the resistance; the 4-volt potential difference means that 4 joules/coul of electrical energy have been lost—in this case converted into heat. By multiplying together the current in amperes and the IR drop in volts, we get the power in watts:

$$0.5 \text{ coul/sec} \times 4 \text{ joules/coul} = 2 \text{ joules/sec} = 2 \text{ watts.}$$

In general,

$$I \text{ (amperes)} \times V \text{ (volts)} = P \text{ (watts).}$$

Of course $P = IV$ is not the only expression possible for determining the power in a circuit or in a part of a circuit. From Ohm's law, $V = IR$; substituting this in the equation given above, we have

$$P = I^2R,$$

which may often be more convenient to use than $P = IV$.

Since power is always the *rate* at which energy is being used or pro-

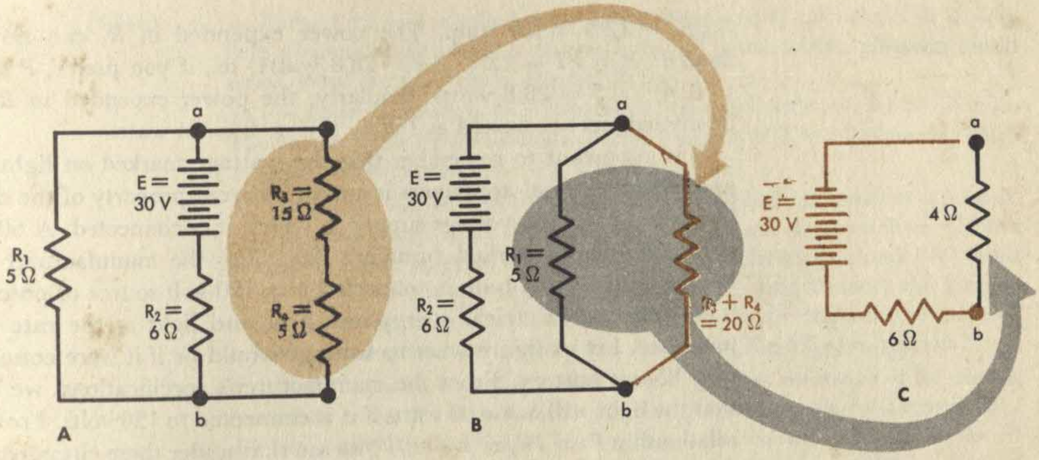


FIG. 13–6 Another example of the simplification of a circuit.

duced or converted into energy of another kind, the total amount of energy converted in some given time is simply the power multiplied by the time:

$$W \text{ (joules)} = P \text{ (watts)} \times t \text{ (seconds)}.$$

In 5 minutes, a 100-watt heater will produce $100 \times 5 \times 60 = 30{,}000$ joules, or 7350 calories of heat.

As an example, let us consider the circuit shown in Fig. 13-6A and calculate the power expended in R_1 and in R_2. If we knew either the potential difference across these resistors, or the current through them, the job would be easy. So we should probably begin by simplifying the circuit to find the total current. The first step is to combine R_3 and R_4 for a total $R_{3+4} = 20$ ohms. This branch $(R_3 + R_4)$ is connected in parallel with R_1. The accidental fact that the drawing shows R_1 on the left side and $R_3 + R_4$ on the right side makes no difference; when the conventional current reaches point b, it must divide in two parts which recombine at a to flow back into the battery. This is shown in Fig. 13-6B. It is apparent that the next step should be to replace the parallel resistances of 5 ohms and 20 ohms with a single equivalent resistance:

$$\frac{1}{R} = \frac{1}{5} + \frac{1}{20} = \frac{1}{4},$$

so $R = 4$ ohms, as shown in Fig. 13-6C. Now, for the whole circuit, there is a total resistance of $6 + 4 = 10$ ohms and an emf of 30 volts, which will result in a current $I = V/R = 30/10 = 3$ amp. The potential difference V_{ab} is just the IR drop through the 4-ohm resistor in Fig. 13-6C, or $V_{ab} = 3 \times 4 = 12$ volts. Armed with this knowledge of V_{ab}, we can turn back to Fig. 13-6B and determine the current through R_1: $I_1 =$

$V_{ab}/R_1 = 12/5 = 2.4$ amp. The power expended in R_1 can now be figured: $P = VI = 12 \times 2.4 = 28.8$ watts; or, if you prefer, $P = I^2R = (2.4)^2 \times 5 = 28.8$ watts. Similarly, the power expended in R_2 can be immediately reckoned as $I^2R = (3)^2 \times 6 = 54$ watts.

It is important to remember that the wattage marked on lightbulbs and other electrical appliances is not an inherent property of the device but depends on the voltage supply to which it is connected. A 60-watt lightbulb is also marked (usually) 120 volts; the manufacturer says, in effect, that *if* the bulb is connected to a 120-volt source of potential, it will convert electrical energy into heat and light at the rate of 60 joules/sec. Let us figure what its wattage would be if it were connected to a 12-volt battery. From the manufacturer's specifications, we know that the light will draw 60 watts if it is connected to 120 volts. From the relationship $P = IV$, or $I = P/V$, we see that under these circumstances the current will be 60/120, or 0.5 amp. Ohm's law, $R = V/I$, shows that the resistance of the filament must be $120/0.5 = 240$ ohms. Assuming that this resistance remains the same, we can see that the current which will flow when the bulb is connected to the 12-volt battery will be $I = V/R = 12/240 = 0.05$ amp. We can now determine the power to be $P = IV = 0.05 \times 12 = 0.6$ watts. The value of $I^2R = (0.05)^2 \times 240 = 0.0025 \times 240 = 0.6$ watts also, to confirm our previous calculation.

Questions

(13-2)

1 A certain resistor is a part of an electric circuit. The potential difference across its terminals is determined by a voltmeter to be 18 volts, and a properly connected ammeter shows that a current of 0.6 amp flows through it. What is the resistance of the resistor?

2 A current of 0.02 amp flows through a resistor in a circuit, and the potential difference between the ends of the resistor is 50 volts. What is its resistance?

3 A uniform wire 1 m long carries a current of 0.1 amp, and has a total resistance of 20 ohms. A voltmeter is connected to the wire at two points: 30 cm from one end, and 40 cm from the other. What potential difference does the voltmeter indicate?

4 A uniform wire 2 m long has a total resistance of 10 ohms. A voltmeter indicates a potential difference of 0.40 volt between point *A*, 50 cm from one end of the wire, and point *B*, 25 cm from the other end. What current flows in the wire?

5 What is the resistance of a silver wire 300 m long, and 2 mm in diameter?

6 What is the resistance of a lead rod 10 cm long and 1 cm in diameter?

(13-3)

7 A cell whose emf is 12.0 volts is in a circuit whose total resistance is 5.0 ohms. What current flows in the circuit?

8 A current of 2.4 amp flows in a circuit containing an emf of 48 volts. What is the total resistance of the circuit?

(13-4) 9. Each of the cells in a battery of 3 cells has an emf of 1.6 volts. This battery is in a circuit which contains resistances of 3, 4, and 5 ohms, all connected in series. What current flows in the circuit?

10. A battery consists of 6 cells in series, each cell with an emf of 1.2 volts. The battery is in a circuit containing resistances of 1, 6, 3, and 8 ohms connected in series. What current flows through the circuit?

11. A battery has an emf of 6 volts and an internal resistance of 1.5 ohms. The external circuit to which its battery is connected contains three 4.5-ohm resistors in series. (a) What is the total resistance of the entire circuit? (b) What is the current in the circuit? (c) What is the *IR* drop through the internal resistance of the battery? (d) What is the terminal voltage of the battery? (e) What is the potential difference across one of the 4.5-ohm resistors?

12. A cell with an emf of 1.5 volts and an internal resistance of 0.1 ohm is connected to four 0.5-ohm resistors in series. (a) What is the total resistance of the entire circuit? (b) What is the current in the circuit? (c) What is the *IR* drop in the internal resistance of the cell? (d) What is the terminal voltage of the cell? (e) What is the potential difference across one of the 0.5-ohm resistors?

13. A 12-volt battery has an internal resistance of 0.5 ohm. What resistance connected to the battery will give it a terminal voltage of 11.0 volts?

14. A 6-volt battery has an internal resistance of 0.05 ohm. What resistance connected across the battery will reduce its terminal voltage to 5.8 volts?

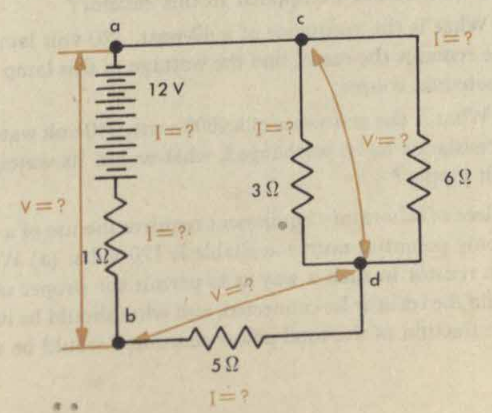
(13-5) 15. Three resistors, 3, 2, and 6 ohms respectively, are connected in parallel. What is the resistance of the single equivalent resistor?

16. Four resistors, two 2-ohm and two 4-ohm, are all connected in parallel. What is the resistance of the single equivalent resistor?

17. A 4-ohm, a 3-ohm, and a 2-ohm resistor are all connected in parallel, and the combination is connected to the terminals of an 8-volt battery of negligible resistance. What current flows in each resistor, and in the battery?

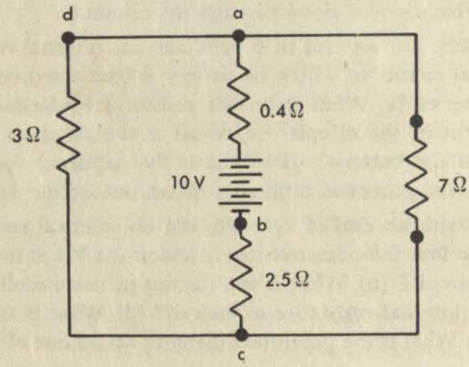
18. Three resistors of 4, 6, and 12 ohms respectively are connected in parallel, and to the terminals of a 10-volt battery whose internal resistance is 0.5 ohm. What current flows in each resistor, and in the battery?

19. Resistances of 3, 5, and 6 ohms are connected as shown to a 12-volt battery whose internal resistance is 1 ohm. (a) What current flows through each



resistance, and through the battery? (b) What is the terminal voltage of the battery (*a-b*)? (c) What is the potential difference measured across *b-d*? (d) Across *c-d*?

20. Resistances of 3, 7, and 2.5 ohms are connected as shown to a 10-volt



battery whose internal resistance is 0.4 ohm. (a) What current flows through each resistance, and through the battery? (b) What is the terminal voltage of the battery (*a-b*)? (c) What is the potential difference across *b-c*? (d) Across *a-c*?

21. A 12-ohm resistor is connected to a 6-volt supply. (a) What is the power dissipated in the resistor? (b) How long will it take to produce 100 joules of heat in the resistor?

22. A resistance of 10^4 ohms is connected to a 500-volt source of potential. (a) At what rate (in watts) is electrical energy converted into heat in the resistance? (b) How long would it take to produce 1000 joules of heat energy?

(13-6) **23.** A 2-ohm resistor can dissipate a maximum of 10 watts without overheating. What is the maximum potential difference that can be applied across the resistor?

24. A 10^6 ohm (1 megohm) resistor is rated at 0.5 watt (i.e., 0.5 watt is the maximum power it can dissipate without overheating). What is the greatest potential difference that should be applied to this resistor?

25. (a) What is the resistance of a 40-watt, 120-volt lamp? (b) Assuming the resistance remains the same, find the wattage of this lamp if it is connected to a 90-volt potential source.

26. (a) What is the resistance of a 4800-watt, 240-volt water heater? (b) Assuming the resistance to be unchanged, what would its wattage be if connected to a 120-volt supply?

27. A piece of laboratory equipment requires the use of a 6-volt, 30-watt lamp, but the only potential source available is 120 volts. (a) Would it be possible to connect a resistor in such a way as to permit the proper use of the lamp? If so, how should the resistor be connected, and what should be its resistance, in ohms? (b) What fraction of the total power consumed would be wasted in heating the resistor?

28. An electromagnet is designed to operate at 30 watts from a 12-volt supply, but the only available source of potential is 28 volts. (a) If it is possible to add an auxiliary resistor to permit the proper operation of the magnet, how should this resistor be connected, and what should its resistance be? (b) What wattage would be dissipated by this auxiliary resistor, and what fraction is this of the power output of the 28-volt supply?

29. A 25-watt, 120-volt lamp and a 100-watt, 120-volt lamp are connected in series across 120 volts. Which lamp will burn the more brightly?

30. Two small 10-gallon, 120-volt water heaters are rated at 500 watts and 1000 watts, respectively. These heaters are connected in series across a 120-volt supply. Which heater will be the first to bring its water to proper temperature?

31. A 4500-watt water heater holds 40 gallons of water at 18°C. (1 gallon = 3.79 kg water.) How long will it take to raise the water temperature to 75°C?

32. An insulated tank holds 2 gallons (see Question 31) of alcohol (density = 0.79 gm/cm^3) at 20°C. The alcohol is heated by submerging a 60-watt lamp in it. How long will it take to raise the temperature to 50°C?

chapter / fourteen

Magnetism

14-1 Magnets and Magnetic Fields

The ancient Chinese knew that slender pieces of certain natural iron ores, when suspended by a string, would assume a definite position with one end pointing approximately north and the other approximately south. It is thus apparent from the behavior of the magnetic compass that the earth is surrounded by a *magnetic field* that affects the orientation of lodestones and other magnets, both natural and artificial. The magnetic field that orients the compass needle manifests itself in many other ways; for example, it deflects the beams of electrically charged particles that come to us from the sun, thus producing the magnificent auroras often visible in the polar regions.

We can use the magnetic field of the earth to magnetize steel rods by holding them in the direction of the magnetic field of the earth and hitting them repeatedly with a hammer. The violent impacts shake the tiny particles of the rod and orient them, at least partially, in the direction of the field. As a matter of fact, all steel objects possess a certain small degree of magnetization induced by the terrestrial magnetic field.

During World War II, much effort was spent to "demagnetize" warships and transports so they would not trigger magnetic mines laid by the enemy.

If we bring two magnetized steel rods close together, we shall find that the "homologous" ends, i.e., the ends that pointed the same way during the magnetization process, repel each other and that if one of the rods is turned around, the ends of the rods will attract one another.

This behavior shows us that a long piece of magnetized material—a splinter of lodestone, a steel bar, or a compass needle—shows its magnetic properties most strongly in regions near its ends, known as the *poles* of the magnet. It also shows that like poles, i.e., poles that point toward the same direction, repel each other and that unlike poles attract. The use of magnets as compasses has given these two sorts of magnetic pole their names: the end that swings toward the geographic north is the north pole of the magnet; the other end, pointing in a southerly direction, is called the south pole of the magnet. It is interesting to notice (in view of the above definitions and in consideration of the fact that unlike magnetic poles attract) that the magnetic pole of the earth located near its geographical north pole is actually its *magnetic south pole*, and vice versa.

The magnetic field around a magnet can be represented by drawing (or imagining) lines, in very much the same way that we represented the electric field in the neighborhood of charged particles. The direction of the magnetic field at any point is shown by the orientation of a small compass needle placed at the point; the magnetic lines of force drawn in this way will be pointed in the same direction as the north pole of the compass needle, as indicated by the arrows in Fig. 14-1.

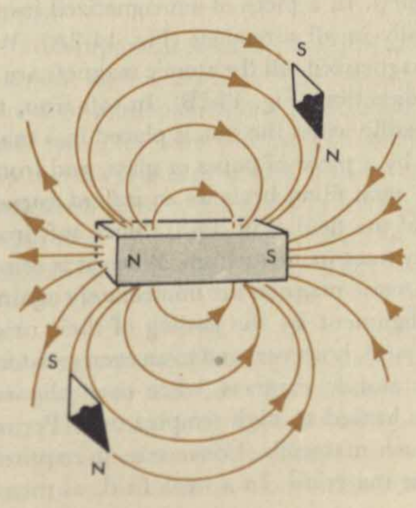


FIG. 14-1 The magnetic field of a bar magnet.

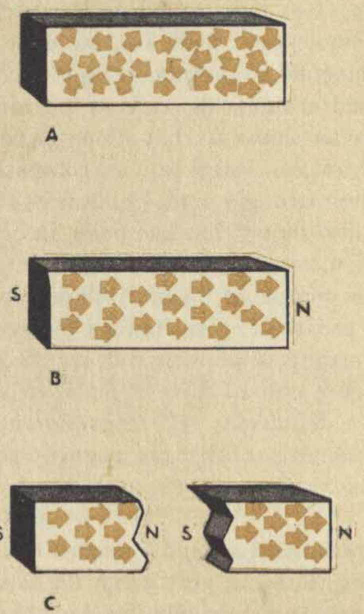


FIG. 14–2 Schematic diagram of alignment of atoms (A) in an unmagnetized bar; (B) in a magnetized bar; (C) in a magnetized bar that has been broken.

Figure 14-2 indicates in a very crude, sketchy manner the idea that magnetic materials are made of atomic particles which are themselves small magnets. In a piece of unmagnetized iron or steel, these particles point equally in all directions (Fig. 14-2A). When the iron or steel is strongly magnetized, all the atomic magnets are lined up in substantially the same direction (Fig. 14-2B). In soft iron, the atomic magnets line up very readily when the iron is placed in a magnetic field. If a magnet is covered by a piece of paper or glass, and iron filings are sprinkled on the paper, each filing becomes an *induced magnet* and aligns itself in the direction of the field (Fig. 14-3). Since soft iron magnetizes readily, it also readily loses its magnetism. When it is removed from the magnetic field, the atomic magnets are immediately again thrown into completely random alignment by the jostling of their neighbors.

In hard steel, however, and to an even greater extent in certain special alloys, the atomic magnets, when once aligned, remain aligned until the steel is heated to high temperatures. Permanent magnets are thus made of such materials. Conversely, it requires a strong field to magnetize these materials. In a weak field, as mentioned above, they must

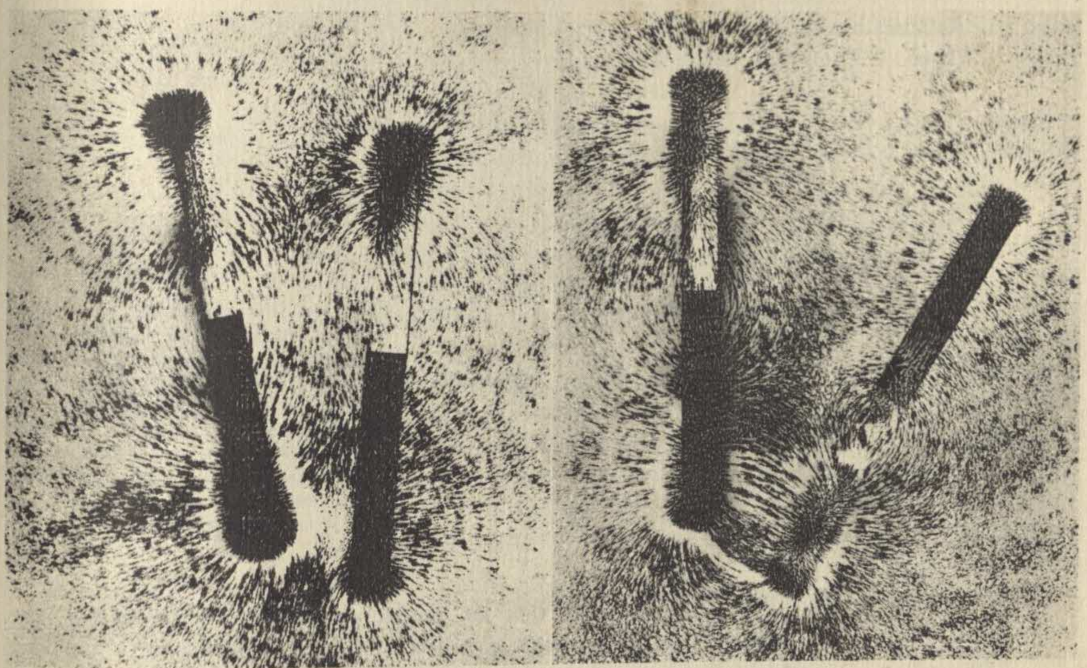


FIG. 14-3 Bar magnets in approximately parallel and antiparallel positions. The resulting magnetic fields are indicated by the alignment of iron filings.

be jarred by hammering to permit the atomic magnets to slip into alignment.

It is important to remember that, unlike positive and negative charges, *magnetic poles must always occur in pairs*, and that it is impossible to cut a north or south pole from a magnet and carry it away. If we cut a magnet into two pieces we will get two smaller magnets, since a new pair of poles will originate at the broken ends. On the basis of the atomic magnetic picture, it is plain to see why north and south poles cannot be separated by breaking a magnet in two. Figure 14-2C indicates the formation of the new pair of poles where the magnet has been broken.

14-2 Currents and Magnetism

The invention of the electric cell and battery, which made it possible to have a continuous flow of electricity, led to the study of various interactions between electric currents and magnets. (*Stationary* electric charges do not interact with magnets at all.) There are several basic laws governing these interactions, all of them discovered early in the nineteenth century. In the year 1820, a Danish physicist, H. C. Oersted

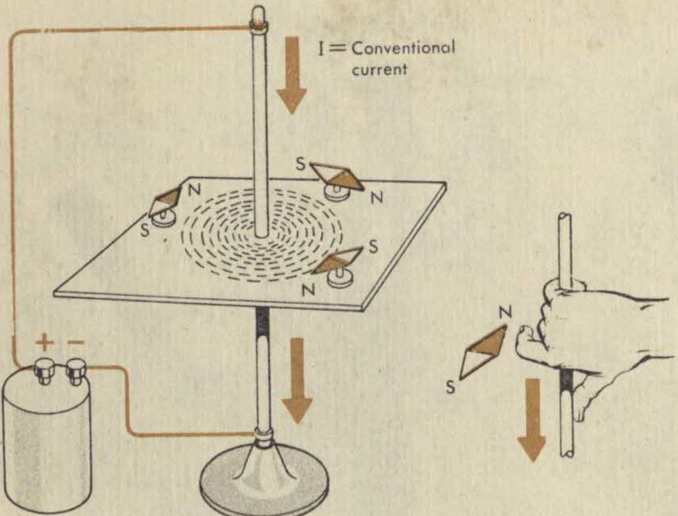
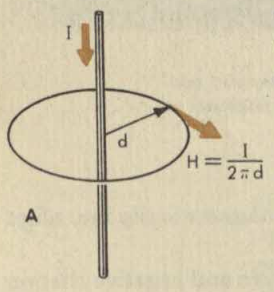
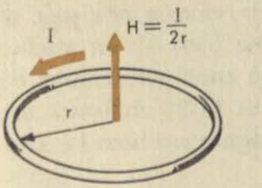


FIG. 14–4 The orientation of magnets (compass needles) in the neighborhood of an electric current. The direction in which the north pole of the needle will point can be found by the following rule: Hold the wire in your **right** hand, so that the thumb points in the direction of the **conventional** current; the fingers then point in the direction of the needle's north pole, and hence also of the magnetic field.



A



B

FIG. 14–5 Magnetic fields produced by electric currents.

(1777–1851), noticed that ***an electric current flowing through a wire deflects a magnetic needle placed in its neighborhood in such a way that the needle assumes a position perpendicular to the plane passing through the wire and through the center of the needle*** (Fig. 14-4). In other words, the magnetic field surrounding a current-carrying wire is in a direction perpendicular to the wire, and the magnetic lines of force representing the field will be concentric circles surrounding the wire. Oersted's discovery was followed up by two French physicists, J. B. Biot (1774–1862) and Félix Savart (1791–1841), who made Oersted's discovery quantitative by finding that ***the strength of the magnetic field (H) created by a current (I) is directly proportional to the strength of the current and inversely proportional to the distance from the wire (d).*** In the system of units we have been using, for a long straight wire (Fig. 14-5A),

$$H = \frac{I \text{ (amperes)}}{2\pi d \text{ (meters)}}.$$

Thus magnetic field strength is seen to be measured in units of amperes/meter.

For a single circular loop of wire of radius r meters (Fig. 14-5B), the field at the center is

$$H = \frac{I \text{ (amperes)}}{2r \text{ (meters)}}.$$

At first glance, it seems that the causes of the magnetic field observed around a bar magnet and of the identical sort of magnetic field around

a current-carrying wire must be entirely different. We can see no current flowing in a bar magnet; it is not connected to a battery or any other source of electric potential, and it can lie on a shelf for decades, quietly maintaining its same field all the while.

Shortly after Oersted's discovery in 1820, André Ampère suggested that in spite of appearances, the causes might actually be the same. He imagined that there might be tiny circulating electric charges within atoms themselves, so that each atom could contain loops of current and therefore itself be a miniature magnet. If this were true, the bar magnet could be explained as the result of the alignment of the atomic electromagnets. This was an inspired guess on Ampère's part, as almost nothing was known about the structure of atoms at that time. We know now that he was fundamentally correct, the only difference being that the field of a bar magnet arises mostly from the spin of its electrons on their axes rather than from their circulation in orbits. But this is a small detail that should not detract from Ampère's credit for the basic idea.

Since the spinning electrons in an atom may be considered to be tiny circulating currents, it is easy to visualize how an atom can actually be a small magnet. This is especially true of some atoms such as those of iron, in which there are more electrons spinning in one direction than in the other.

14-3 Force on a Moving Charge

We mentioned before that a magnet and a *stationary* electric charge do not have any effect on each other. A charged pith ball and a strongly magnetized compass needle will completely ignore one another. An electric current, however, is nothing more than a stream of *moving* charges; and we have seen that these moving charges create a magnetic field whose direction is at right angles to the direction of the current. It should therefore not be too surprising to find that a charge which is *moving* in a magnetic field experiences a force, and that ***the force is at right angles to both the velocity of the charge and the direction of the field.***

Figure 14-6 shows how the direction of the force on the moving charge can be predicted. It shows a *right* hand held flat, the thumb extended at right angles to the other fingers. Orient the hand so the fingers point in the direction of the magnetic field; now turn the hand (keeping the fingers still pointed in the field direction) so the thumb points in the direction the *positive* charge is moving. (If the moving charge is negative, one needs only point the thumb in a direction *opposite* its velocity—a negative charge moving, say, to the left is exactly equivalent to a positive charge moving to the right.) The direction in which you would now push forward with the palm of your hand is the direction in which the magnetic field pushes against the moving charge.

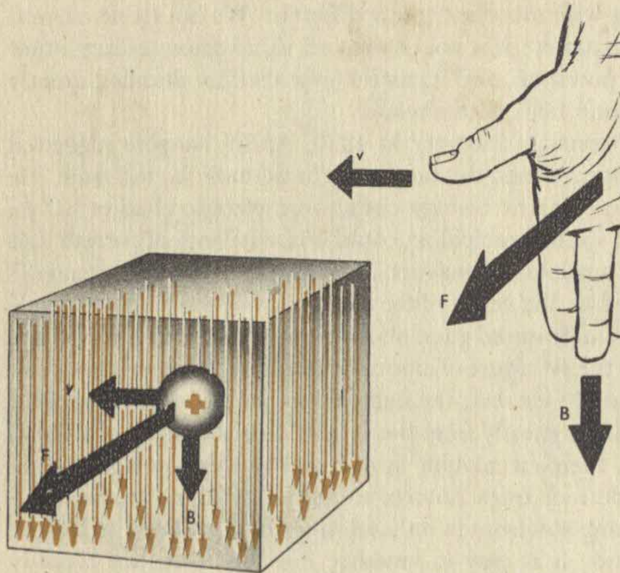


FIG. 14-6 Force on a charged particle moving in a magnetic field.

The magnitude of this force is

$$F = qv\mu H$$

with F measured in newtons; q in coulombs; v, the component of the velocity perpendicular to the field, in meters/second; and H in amperes/meter.

The symbol μ (the Greek letter *mu*) needs some explaining. In the equation above, we have arbitrarily chosen the units of F, q, v, and H to be those with which we were already familiar in many other applications. In order to reconcile all these units, a proper numerical constant is needed. It is as though we had decided to state Newton's second law with force measured in pounds, mass in grams, and acceleration in meters/second^2. If we want to use these units, we cannot write $F = ma$ but must insert the proper conversion factor so that the equation reads $F = 2.24 \times 10^{-4}\ ma$.

The factor μ is called the *permeability* of the material in which the magnetic field is located. In a vacuum (or in air, which for most practical purposes is magnetically the same),

$$\mu = 4\pi \times 10^{-7} = 12.57 \times 10^{-7}.$$

The product μH finds much more use than does H alone; and because it appears so frequently, it is given its own private symbol: $\mu H = B$, which has the name *magnetic flux density* or *magnetic induction*, and is measured in units of *webers/square meter*, which we shall discuss in the next section. Using this symbol B, the expression for the force on a charged particle moving in a magnetic field can be written in the following simpler form:

$$F = Bqv.$$

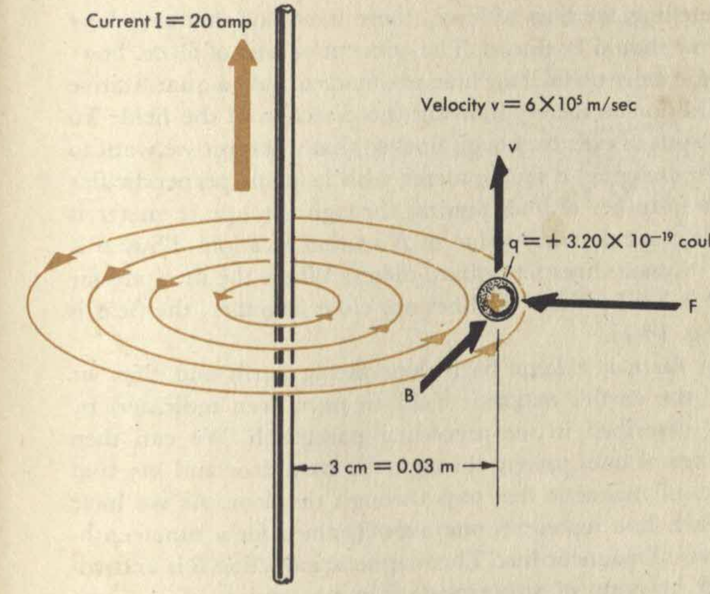


FIG. 14-7 Force on an alpha particle moving in the magnetic field of a current-carrying conductor.

As an example of the force exerted on a moving charge, imagine an α*(alpha) particle* (charge $+3.20 \times 10^{-19}$ coul) 3.0 cm from a long, straight wire, traveling with a velocity of 6×10^5 m/sec parallel to the wire. The wire carries a current of 20 amp, which flows in the same direction in which the α particle is traveling. What is the force on the particle? (See Fig. 14-7.) We can first find the magnetic field H at a distance of 3 cm, or 0.03 m, from the wire:

$$H = \frac{I}{2\pi d} = \frac{20}{2\pi \times 0.03} = 106 \text{ amp/m}.$$

Since the field is in air, the magnetic induction B is $12.57 \times 10^{-7} \times H = 1.33 \times 10^{-4}$ weber/m². It will not be necessary to calculate the component of the velocity perpendicular to the field, because in this problem the velocity itself *is* perpendicular to the field, and we get $F = qvB = 3.20 \times 10^{-19} \times 6 \times 10^5 \times 1.33 \times 10^{-4} = 2.55 \times 10^{-17}$ nt. The direction of the force (use Fig. 14-6) is directly in toward the wire. Since this hand rule is designed to use the velocity of a *positive* charge, we would get an opposite answer for a negatively charged particle moving in the same direction.

14-4 Magnetic Flux

In the preceding section, B has been measured in *webers/square meter*, which serves to introduce the concept of *magnetic flux*. If we want only to show the direction of the field around a magnet or a system of electric

currents by sketching the lines of force, there is no law that says how many or how few should be drawn. The concept of lines of force, however, can be made more useful if the lines are made to have a quantitative meaning, in addition to merely showing the direction of the field. To do this, we may put in exactly enough lines so that wherever we want to draw (or, better, imagine) a square meter with its plane perpendicular to the field, the number of lines passing through the square meter is exactly equal to the numerical value of B at that location. Thus B is represented as so many lines per square meter. Where the lines are far apart, the field is weak; where the lines are closer together, the field is stronger (see Fig. 14-1).

Suppose that there is a large barn door facing north and that we have imagined the earth's magnetic field to have been indicated by lines drawn as described in the preceding paragraph. We can then count the number of lines passing through the barn door and say that so many webers of magnetic flux pass through the door. As we have drawn them, each line represents one *weber* (named for a nineteenth-century physicist) of magnetic flux. The magnetic induction B is accordingly measured in units of webers/square meter and is often quite logically called the *flux density*. The symbol Φ (the Greek capital letter *phi*) is generally used to represent flux. If we have some area A (for simplicity taken perpendicular to the field) through which a total flux of Φ passes, then the average flux density within the area is $B = \Phi/A$ webers/m^2. Conversely, an area A perpendicular to a given magnetic induction B is cut by $\Phi = AB$ webers of flux.

14-5 Solenoids and Electromagnets

If a current-carrying wire is wound in the form of a helix (Fig. 14-8A), it is called a *solenoid* and can be considered as a large number N of single turns of wire connected in series. It can be shown (with calculus, again!) that the field within a solenoid is given by

$$H = \frac{NI}{l}$$

and the flux density or magnetic induction is, of course,

$$B = \mu H = \frac{\mu NI}{l}.$$

The flux density can be increased by increasing the current I or by increasing N/l, the number of turns per meter, or by increasing μ.

It is easy to see how one might change the current in a coil or rewind it to change its N/l; but how does one go about changing μ? The permeability factor μ not only takes into account the conversion of units but also the effect of the material in which the field is established. Suppose, for example, a cylinder of iron were slipped inside the coils of the sole-

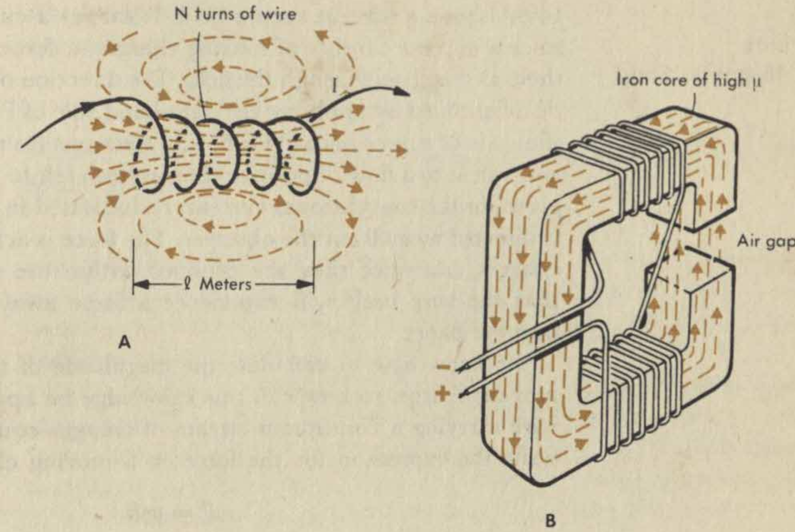


FIG. 14-8 A simple solenoid, and one form of electromagnet.

noid in Fig. 14-8A. When the current is turned on, the field established by the current will at least partially align the atomic magnets of the iron to point with the field, so that each one contributes a little extra flux in the same direction. The end result of all these atomic contributions is that we may expect to have more flux, and a higher flux density B with the iron core than without it. The *relative permeability*—that is, $\mu_{\text{material}}/\mu_{\text{vacuum}}$—can be as high as several thousand for some iron alloys and several hundred thousand for other specially manufactured magnetic materials.

Such materials, which markedly increase the flux density, are called *ferromagnetic*; this group, besides iron and many special alloys, also includes the elements cobalt and nickel. For most materials, however, the relative permeability is very nearly 1. Some (like bismuth, relative permeability 0.99983) are *diamagnetic*, and make the flux weaker; others (aluminum, relative permeability 1.00002) are *paramagnetic*, and contribute very weakly.

Electromagnets are nearly always wound on a ferromagnetic core. With many turns of wire, which are sometimes water-cooled to permit them to carry a high current without overheating, a high flux density can be established. The electromagnet in Fig. 14-8B is merely a solenoid bent into a C, its winding being in two parts.

Suppose we place a current-carrying wire in a strong magnetic field, such as the air gap of the electromagnet shown in Fig. 14-8B. Figure

14-6 Currents in a Magnetic Field

14-9A shows a wire in such a field. It carries a current as shown; and since a current consists of moving charges, a force must be exerted on them as they pass through the field. The direction of this force can easily be determined by applying the right-hand rule of Fig. 14-6. The current shown is of course actually a flow of electrons from right to left; this is equivalent to a flow of positive charges from left to right in the direction given for the conventional current. As indicated in Fig. 14-9B, the force is directed away from the observer. The force is actually on the moving charges, but since they are confined within the wire, we can expect that the wire itself will experience a force away from the observer, into the paper.

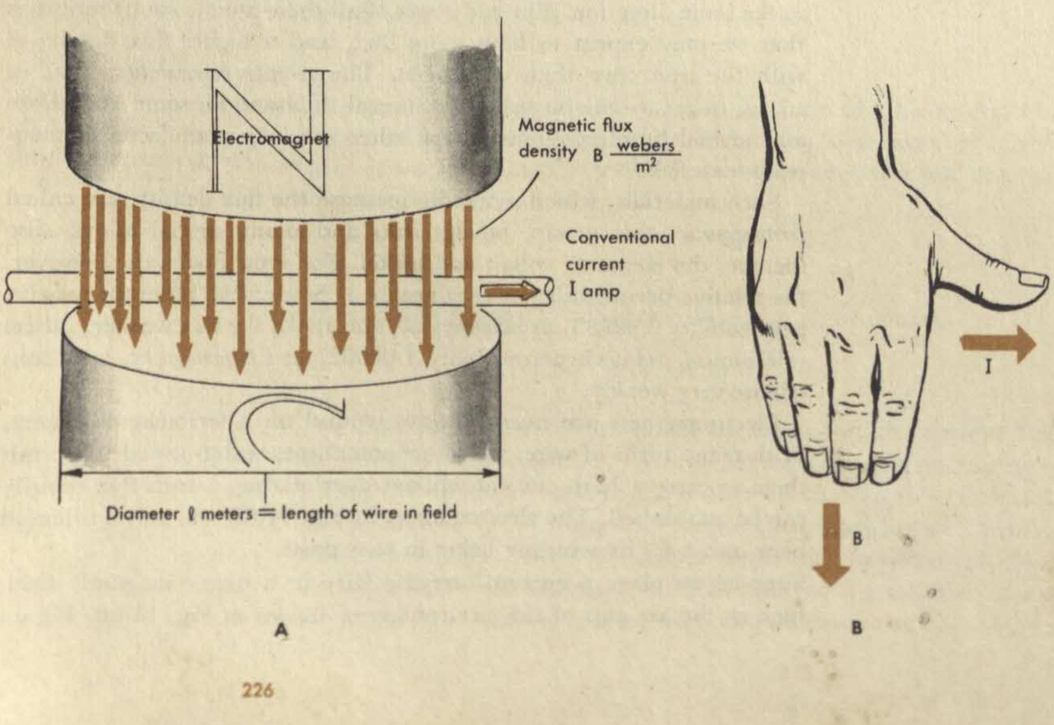
We know how to calculate the magnitude of the force on a single moving charge, so how can this knowledge be applied to the case of a wire carrying a continuous stream of charges equal to I amp? Let us write the expression for the force on a moving charge:

$$F = qvB.$$

Dimensionally, this equation is

$$\text{Newtons} = \frac{\text{coulombs} \times \text{meters} \times \text{webers}}{\text{second} \times \text{meter}^2}.$$

FIG. 14–9 A current-carrying wire in a magnetic field.



Absolutely nothing is changed if we quietly move the "second" in the denominator until it is under the "coulombs":

$$\text{Newtons} = \frac{\text{coulombs} \times \text{meters} \times \text{webers}}{\text{second} \quad \times \quad \text{meter}^2}.$$

Since a coulomb/second is an ampere, which measures current, the equation becomes

$$F = IlB$$

where l is the length of wire in the magnetic field. In the original equation, v was the component of velocity perpendicular to the field; in the equation above, l must therefore be interpreted as the component of the length perpendicular to the field.

Let us make a numerical example of Fig. 14-9: the diameter of the magnet poles is 20 cm, and the flux density between them is 1.65 webers/m². The wire passes through the center of the field (which means that 20 cm, or 0.20 m, of the wire is in the field) and is perpendicular to it. The current is 30 amp. For the force exerted on the wire,

$$F = IlB = 30 \times 0.20 \times 1.65 = 9.90 \text{ nt}.$$

(We have already determined that the force will be into the paper.)

14-7 Galvanometer, Voltmeter, Ammeter

One way in which the above force is utilized is in the construction of electric meters of many kinds. The *galvanometer*, besides being an important instrument itself, is the actual working element in most voltmeters and ammeters. There are many kinds of galvanometers, which

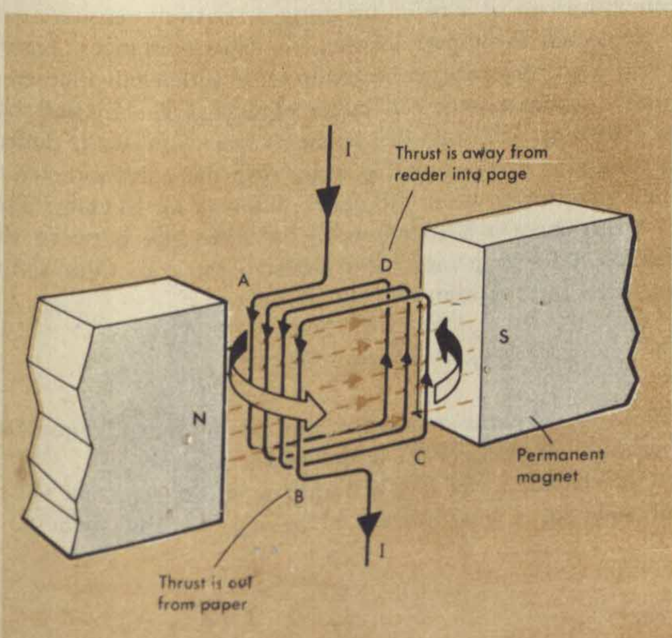


FIG. 14-10 A moving-coil, or D'Arsonval type, galvanometer.

are designed to measure very small currents and potential differences, but the most common is the d'Arsonval type, in which a very light coil of wire is pivoted in a strong magnetic field (Fig. 14-10).

The direction of the magnetic field, which is defined as the direction in which the north pole of a compass needle would point, always must run from the north pole of a magnet to the south pole, which, in Fig. 14-10, is from left to right. Applying the right-hand rule, we see that the wires of the coil from A to B experience a force toward the reader, out from the plane of the page; the wires from C to D are thrust away from the reader into the page. These forces cause a torque that rotates the coil about its central axis I-I against the restraining torque of a small spring not shown in the drawing. The rotating torque is proportional to the current in the coil; and hence the greater the current, the greater the angle through which the coil is rotated. In many meters, an indicating needle is attached directly to the coil, but for very sensitive galvanometers a beam of light reflected from a tiny mirror is used instead. Some d'Arsonval galvanometers can indicate currents as small as 10^{-10} amp. (There are other more complicated types of current-measuring devices which will give a reading for 10^{-17} amp!)

Even the coarsest, most insensitive galvanometer is not adapted to measure large currents or high voltages. If it were so designed, the galvanometer coil itself would add too much resistance when put in series in a circuit as an ammeter and would change the very current it was intended to measure. On the other hand, if it were used directly as a voltmeter, the coil resistance would be so low that considerable current would flow through the galvanometer, thereby changing the IR drop it was supposed to measure accurately.

These difficulties can be avoided by using a relatively sensitive galvanometer connected to properly chosen auxiliary resistances. Figure 14-11A shows a high resistance connected in series with a galvanometer. Let us say that the galvanometer coil has a resistance of 50 ohms and that a current of 0.01 amp will cause the needle to make a full-scale deflection. What resistance must we put in series with this galvanometer to convert it into a voltmeter which will deflect full scale for 10 volts? This means that when there is a 10-volt potential difference between the meter terminals, 0.01 amp must flow through the coil. Ohm's law gives the required total resistance of the voltmeter:

$$R_T = \frac{V}{I} = \frac{10}{0.01} = 1000 \text{ ohms.}$$

Of this 1000 ohms, 50 ohms are already supplied by the galvanometer coil itself, leaving 950 ohms for the added series resistance.

To make the galvanometer into a 5-amp ammeter, i.e., one which deflects full scale for a total current of 5 amp, we must provide a

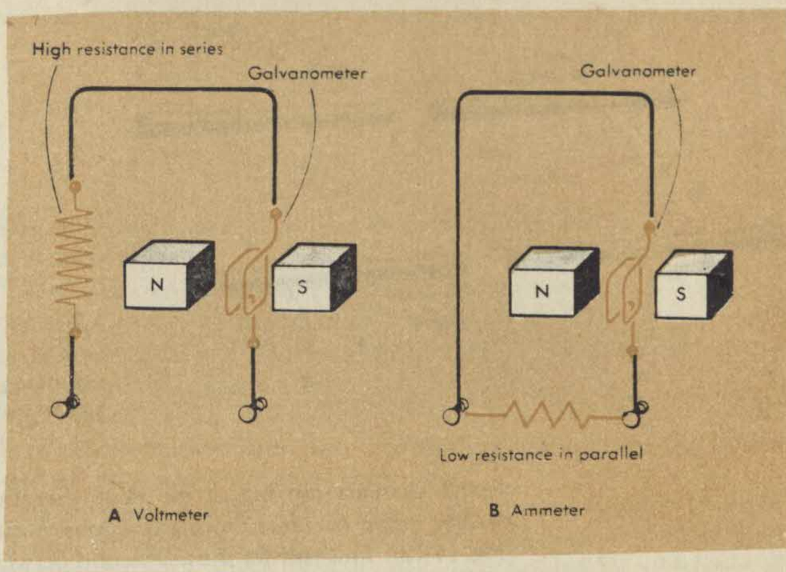


FIG. 14–11 The auxiliary resistances used with a galvanometer to make a voltmeter or an ammeter.

shunt, or low-resistance path in parallel with the galvanometer (Fig. 14-11B). Of the total 5 amp, 0.01 amp must flow through the galvanometer and the remaining 4.99 amp must go through the shunt. Since the two parallel paths are connected to the same terminals, the potential difference (which is the *IR* drop) across each is the same, and we can say

$$I_S R_S = I_G R_G$$

or

$$4.99 R_S = 0.01 \times 50$$
$$R_S = 0.1002 \text{ ohm.}$$

All the common types of electric motor also operate on this same principle—that a current-carrying conductor in a magnetic field receives a thrust perpendicular to both the field and the current in the direction given by the right-hand rule. In all motors except some special kinds and some toy motors, the magnetic fields are provided by electromagnets.

14-8 Interactions between Currents

So far, we have been talking about the interaction between currents and magnetic fields. What about the interaction between two currents? This was first studied by Ampère, who showed that two wires carrying

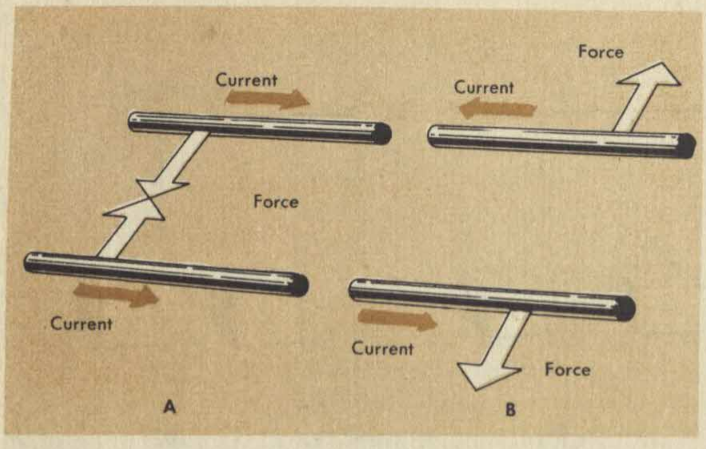


FIG. 14–12 The forces acting on two parallel current-carrying wires.

electric currents flowing in the same direction are attracted to one another, while currents flowing in opposite directions repel each other (Fig. 14-12). This rule for the interaction between currents can be shown to follow from the behavior of current-carrying conductors in magnetic fields. Each wire is in the magnetic field of the other wire; each experiences a force which pulls them apart or pushes them together, depending on the direction of their currents.

Take as an example two long, straight parallel wires M and N, separated by a distance of a meters (Fig. 14-13). The magnetic field intensity H caused by the current I_M is $I_M/2\pi a$ at a distance a, where I_N is in the field. The flux density B is μH, or $12.57 \times 10^{-7} \times I_M/2\pi a$, if M and N are separated by air or vacuum. We cannot determine the total resultant

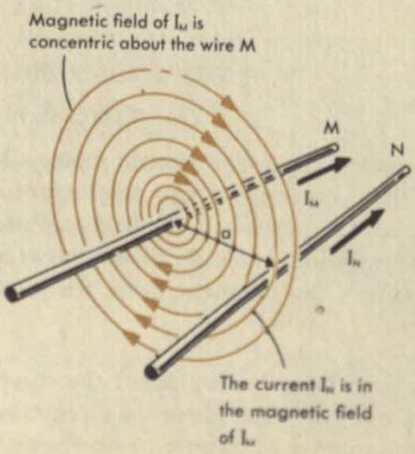


FIG. 14–13 A current-carrying wire in the magnetic field of another parallel current-carrying wire.

force on N, because it is indefinitely long, but for a 1-m length of N, the force given by

$$F = BI_N l$$

becomes

$$F = 12.57 \times 10^{-7} \times \frac{I_M I_N}{2\pi a}.$$

The right-hand rule shows that this force on N is in a direction to pull it in toward M.

From Newton's third law, we know in advance that if M exerts the above force on N by means of M's magnetic field, then N must exert an exactly equal and opposite force on M. A short calculation will convince you that this is indeed true.

If the two currents flow in opposite directions, the magnitude of the force will remain the same, but the right-hand rule tells us that in this case the conductors will repel each other.

As an example of the application of the formula worked out above (known as *Ampère's law*), let us consider two parallel wires 50 cm long in a large electric motor. These wires are separated by a distance of 1 cm, and the insulation between them, being almost perfectly non-magnetic, has the same permeability as air. A short circuit develops in the motor, with the result that a sudden current of 5000 amp flows through the two wires (a short circuit is the formation of a very low resistance path for current, generally due to mechanical damage to insulation). What force will tend to pull the wires apart, assuming that the currents run in opposite directions? A direct substitution of numerical values in Ampère's law gives

$$\begin{aligned} F &= 12.57 \times 10^{-7} \times \frac{I_1 I_2}{2\pi a} \\ &= \frac{12.57 \times 10^{-7} \times 5000 \times 5000}{2\pi \times 0.01} \\ &= 500 \text{ nt/m}. \end{aligned}$$

Since the length of the wires is 50 cm, or 0.5 m, the force on each wire will be 250 nt. A newton is equal to a force of about a quarter of a pound, so the wires will be shoved apart by a force of more than 60 lb.

14-9 Generation of Electric Currents

We have seen that if a wire carrying current is placed in a magnetic field, a force acts on the wire and will move it if it is not fastened in place. In 1831, it occurred to the great British experimental physicist and chemist Michael Faraday, and independently to the American physicist Joseph Henry, that a converse effect should also be observable. That is,

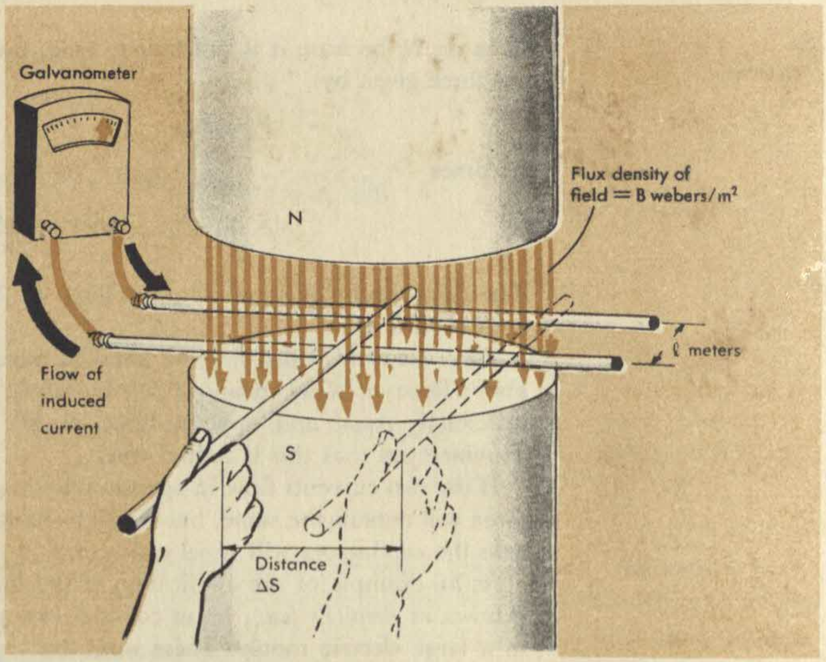


FIG. 14–14 Inducing a current in an electric circuit.

if a conductor is moved in a magnetic field, a current should be generated in the conductor. Both men observed that this actually did happen.

Let us arrange an apparatus as shown in Fig. 14-14. It consists of a pair of parallel bare conductors spaced l meters apart in a magnetic field of flux density B. Across these is placed another bare conductor making contact with the two parallel wires. A galvanometer is connected to the parallels to complete an electrical circuit and to indicate any current that flows. Now, if the crosswire is moved to the left, the galvanometer needle will deflect in one direction; when the crosswire is moved to the right, the galvanometer deflects in the opposite direction.

Suppose the crosswire is moved to the left a distance of Δs meters in a time Δt; the galvanometer will show us that a current flows in the circuit. A current does not flow unless there is an emf in the circuit; so apparently a potential difference e has been created in the moving wire. Any electric current expends energy while flowing, and the principle of conservation of energy tells us that this electric energy must be supplied by the work we do in moving the crosswire. Assume (ignoring friction and the inertia of the wire) that a force F is required to move the wire; the work done in moving it a distance Δs is equal to $F\Delta s$.

The force F is evidently needed because there is a current I flowing through the l meters of the crosswire, and the crosswire is in a magnetic

field. In the previous section, we found that a force $F = BIl$ would be exerted on the wire; this is the opposing force we must overcome to move the wire, and thus the work done can be expressed as

$$W = F\Delta s = BIl\Delta s.$$

In Fig. 14-14, there is a complete electrical circuit which encloses a certain amount of magnetic flux. Now as the crosswire is moved to the left, the area of the circuit is reduced and it accordingly encloses less flux. The circuit area is reduced by $l\Delta s$ m^2, and as a result it encloses $Bl\Delta s$ webers less flux than it did; that is, $\Delta\Phi = Bl\Delta s$ webers.

Thus the work done in moving the wire is

$$W = BIl\Delta s = I\Delta\Phi.$$

All that remains to be done now is to divide both sides of this equation by the time Δt:

$$\frac{W}{\Delta t} = \frac{I\Delta\Phi}{\Delta t}.$$

We recognize $W/\Delta t$, work per unit time, as being power, which in an electrical circuit equals the potential times the current. The substitution of eI for $W/\Delta t$ gives

$$eI = \frac{I\Delta\Phi}{\Delta t}$$

or

$$e = \frac{\Delta\Phi}{\Delta t}.$$

This equation gives the relationship we have been working toward. It says that whenever the magnetic flux enclosed by a circuit changes, a potential equal to the rate of change of the flux is generated. (Assuming, naturally, that our units are correct, potential will be in volts and rate of change of flux in webers/second.)

A rearrangement gives us another statement that is often useful. Power is also given as force times velocity, that is,

$$P = Fv$$

or

$$Ie = BIlv$$

from which

$$e = Blv.$$

The direction of the emf and the resulting current can be determined by a special application of the rule of Fig. 14-6. In Fig. 14-14, when the

crosswire is moved to the left, the hypothetical + charges of the "conventional" notation are forcibly carried along with the rest of the wire thus constituting a current to the left. The right-hand rule of Fig. 14-6 shows that these charges will experience a force driving them out of the paper toward the hand holding the crosswire, thus driving the current in a clockwise circuit through the galvanometer.

In science, there is a very general principle known as *the principle of Le Chatelier*. It states that ***whenever we undertake any action to change an existing physical system, the system reacts in such a way as to oppose our action.*** Reflection will show that this principle necessarily follows if we are to have conservation of energy; otherwise perpetual-motion machines would be a dime a dozen, and we would be able to create unlimited amounts of energy from any small starting push. Le Chatelier's principle, when applied to the interactions of currents and magnetic fields, is known as *Lenz's law*. Applying Lenz's law to Fig. 14-14, we see that the induced current in the crosswire must flow in such a direction that its reaction with the magnetic field will produce a force acting to the right, opposing the motion of the wire. Our rule shows that the current must flow through the crosswire toward the hand, in the same direction that we had already concluded from a different line of reasoning.

Figure 14-15A shows a coil of 100 turns of wire wound on a circular frame 20 cm in diameter. The plane of the coil is parallel to the earth's magnetic field, which has a flux density of 0.5 gauss. (The *gauss* is a commonly used flux density unit deriving from the CGS system, in which flux is measured in *maxwells* rather than in webers. A gauss is 1 maxwell/centimeter^2 and equals 10^{-4} weber/meter^2.) The coil is pivoted on an axis, OO, perpendicular to the field. If the coil is rotated through 90° about OO in 0.2 sec, what average voltage will be generated? As the coil is first oriented, no flux passes through it; after 90° rotation, the flux will pass through the entire area within the coil. This area is $\pi r^2 = 0.0314$ m^2, and the flux density of the field is 0.5 gauss, or 5×10^{-5} weber/m^2. Therefore, for each turn of the coil, the enclosed flux will change from zero to $0.0314 \times 5 \times 10^{-5} = 1.57 \times 10^{-6}$ weber. This change takes place in 0.2 sec, so the induced emf will be $\Delta\Phi/\Delta t = 1.57 \times 10^{-6}/0.2 = 7.85 \times 10^{-6}$ v. This emf is induced in each one of the 100 turns, and hence the total average voltage will be 7.85×10^{-4} volt, or 0.785 millivolt. A *millivolt* is, like all "milli's," a *thousandth* of a volt.

Another example is given in Fig. 14-15B. A single wire AB is drawn through the air gap of a magnet at a speed of 1 cm/sec. The magnet poles measure 5 cm × 5 cm, and the voltage induced while AB is passing between the poles is 0.05 millivolt. To find the flux density in the air gap, it will be convenient to use $e = Blv$, which gives us

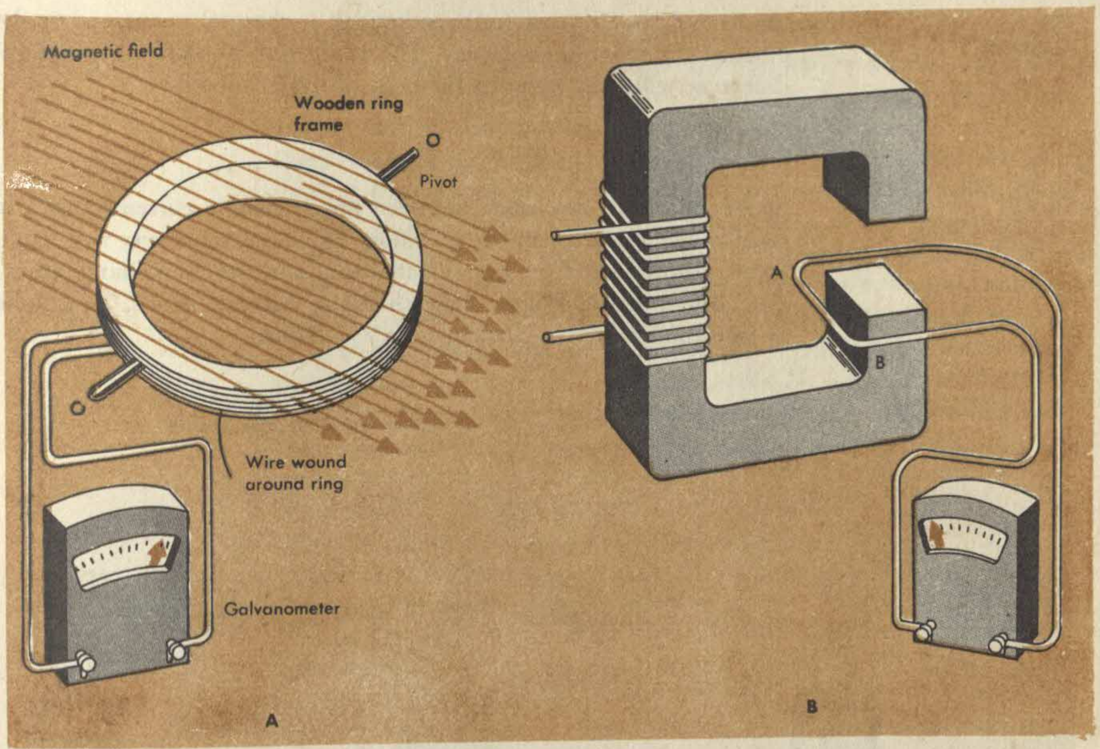


FIG. 14–15 Examples of emf induced in conductors moving in a magnetic field.

$$5 \times 10^{-5} = B \times 5 \times 10^{-2} \times 10^{-2}$$

or

$$B = 0.1 \text{ weber/m}^2$$
$$= 1000 \text{ gauss.}$$

(In checking the above equation against the given data, remember that everything had to be converted from centimeters to meters.)

In deriving the relationships

$$e = \frac{\Delta\Phi}{\Delta t} \quad \text{and} \quad e = Blv$$

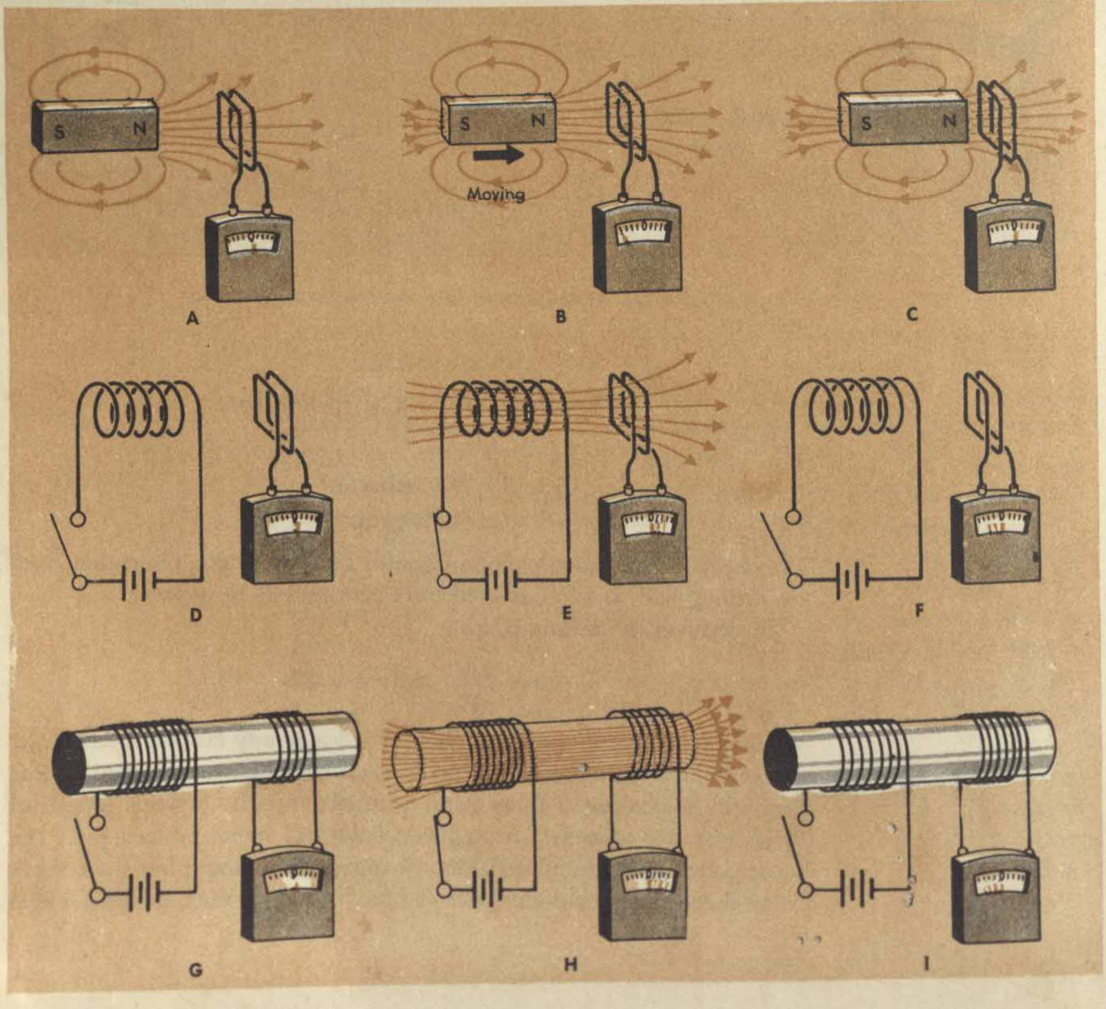
we used a complete circuit through which a steady current could flow. However, the formulas for induced potential which we derived in this way still hold, even if there is no complete circuit. A single length of conductor moved across a magnetic field will have induced in it the same potential as calculated from the equations above; but since there is no closed path through which it can drive a current, the only effect

can be to drive electrons (or hypothetical + charges) to one end or the other of the moving conductor. The result will be that this potential difference will exist between the ends of the conductor.

14-10 Changing Flux

So far, all of our discussion has been about electric currents that run steadily and smoothly in one direction. However, some very important consequences come from currents that stop and start and change direction. Before we investigate these changing currents, it will be helpful to take another look at the equation $e = \Delta\Phi/\Delta t$.

FIG. 14-16 An electric potential is induced when the flux passing through a coil is changing.



Although this equation was derived by considering a conductor moving in a magnetic field, the equation holds for *any* situation in which the flux enclosed by a circuit changes. Figure 14-16 shows examples. In Fig. 14-16A, the coil connected to the galvanometer is cut by very little flux; while the magnet is being moved toward it (Fig. 14-16B), the flux passing through the coil is increasing, and the needle is deflected to the left. In Fig. 14-16C, the magnet is stationary, and although considerable flux threads the coil, it is no longer changing, and the galvanometer reads zero. If the magnet were withdrawn, the flux would reduce, and the galvanometer would deflect to the right. If this whole experiment were repeated, with the magnet turned so its south pole were nearer the coil, the galvanometer deflections would be reversed in direction.

In Fig. 14-16D, the galvanometer coil is placed near another coil connected to a battery, but the switch is open and there is no current and no flux. In Fig. 14-16E, the switch has been closed; in a tiny fraction of a second, the current and the flux increase to their maximum. During this brief period of growing flux, the needle swings over and immediately returns to zero when the flux has reached its maximum and stops changing. The switch has been opened again in Fig. 14-16F, the current and flux drop to zero, and as the flux rapidly decreases, the needle takes a brief swing, this time to the left. Lenz's law tells us that in the case of the *increasing* flux (Fig. 14-16D to E), the induced current will flow in the direction that will make a flux opposed to that of the battery-connected coil, in order to oppose its increase. When the switch is opened, the induced current will flow in the direction that will produce flux in the same direction as the flux of the battery-connected coil, in order to keep it from decreasing.

FIG. 14–17 Primary and secondary windings of a simple transformer.

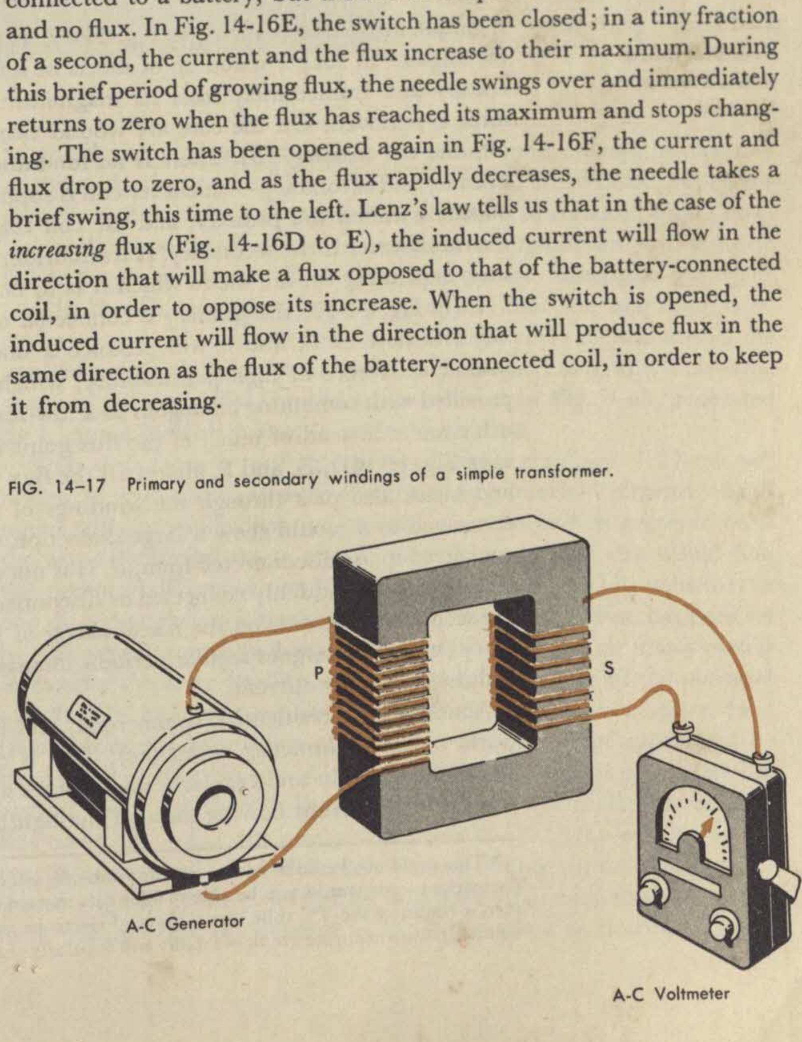


Figure 14-16G, H, and I repeat the procedure of D, E, and F, except that both coils are wound on a piece of soft iron. The permeability of the iron μ is much greater than that of air, so the induced flux is greater, and the galvanometer needle takes a larger swing.

All the discussion above seems a long way from the generation and distribution of the alternating current that is used in nearly all our homes, offices, and factories. It does, however, emphasize the principle that whenever the flux within a coil changes, an emf is induced in the coil, and a current will flow if the coil is a part of a complete circuit. It is this principle that makes the transformer possible, and without the transformer, our present wide distribution of cheap electric power would not be possible.

14-11 Transformers and Alternating Current

A simple *transformer* is shown in Fig. 14-17. It contains a *core*, generally rectangular in shape, made up of sheets of soft iron which magnetizes readily and immediately loses its magnetism when the magnetizing force is removed. On the core are wound two coils, called the *primary* (P) and the *secondary* (S). There is no fundamental difference between primary and secondary—whichever coil is supplied with power from the outside is called the primary, and the coil that supplies the output of the transformer is the secondary. A transformer can be connected either way, depending on the use to which it is to be put.

If (in Fig. 14-17) the primary is connected to a battery, current will begin to flow in P, and it will produce a magnetic flux. The coil is wound on a core of high permeability, which means that the flux is provided with something analogous to a low-resistance path for an electric circuit. Instead of much of the flux going into the surrounding air, as in Fig. 14-16D, E, and F, almost all the flux will stay within the core and hence also pass through the windings of the secondary. A meter connected to S would show a large deflection whenever a battery was connected to, or disconnected from, P. It is not necessary, however, that the primary be suddenly connected or disconnected; the emf induced in the secondary depends on the *rate of change* of the flux, and if the primary current undergoes regular periodic changes, so will the flux, and so will the secondary current.

Almost all the residential, commercial, and industrial current in the world today is *alternating current* (ac). This is current that reverses its direction smoothly and regularly many times a second. If you were to analyze* the current flowing through the lightbulb in your desk lamp,

* This could not be done with a regular ammeter, which could not follow the rapid fluctuations—nor would you be able to follow its motion if it did. An *oscilloscope*, however, a cousin of the TV tube, has a beam of electrons as its only moving part and is able to draw an accurate graph of rapidly and regularly changing currents or potentials.

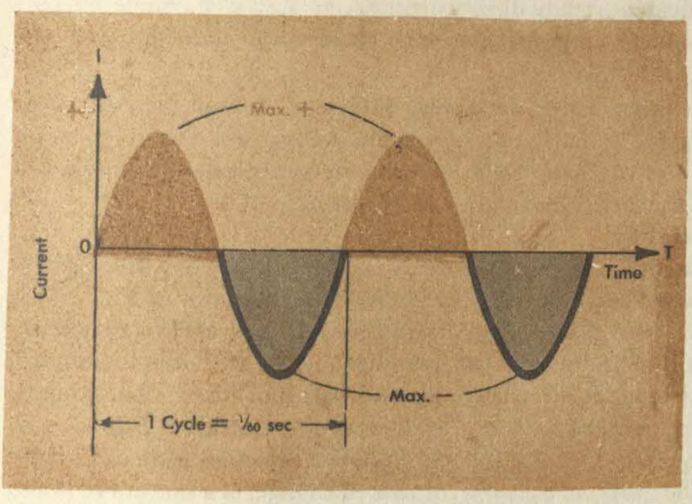


FIG. 14-18 Graph of the changes in an alternating current.

it would be found to fluctuate in magnitude and change in direction as shown in the graph of Fig. 14-18. Starting at some instant when the current is zero, it rises to a maximum, then falls to zero and begins flowing in the opposite direction, to again reach a maximum and again to fall to zero—and so on. These described changes constitute one cycle, and in the United States require $\frac{1}{60}$ sec, so that the frequency is 60 cycles/second. The smooth rise and fall of the current (and also of the applied potential) follows a *sinusoidal*, or *sine curve*. The graph is the same shape as the sinusoidal transverse wave shown in Fig. 8-4A, generated by a source moving in simple harmonic motion.

Your desk lamp is probably plugged into a standard "120-volt ac" outlet. If you were to analyze the changing potential difference of this outlet with an oscilloscope, you would of course find its graph to be of the same sinusoidal shape as the current graph. But you would find the potential difference would rise to a maximum of 170 volts first in one direction and then in the other. This maximum value has been set by the power company because it results (over one or many cycles) in exactly the same heating effect as though a steady direct potential of 120 volts had been applied to a lamp or an iron or a heater. In a circuit carrying alternating current, the power will of course not be steady. When the current is I_{max}, electrical energy will be converted into heat at the rate of $I_{max}^2 R$; when I is zero, the power will likewise be zero. But when $I = -I_{max}$, the power will be the same $I_{max}^2 R$—which way the current flows makes no difference to an electric iron. Figure 14-19 shows a graph of the power of an ac circuit. It is easy to show with a little calculus that the average power is the same as that in a circuit

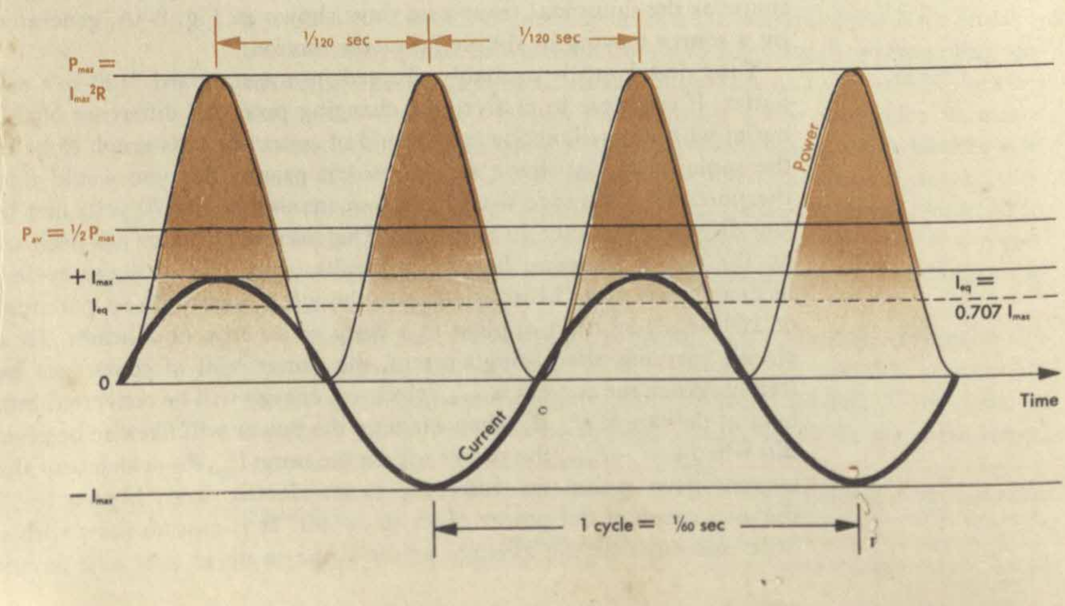
with an equivalent steady direct current $I_{eq} = I_{max}/\sqrt{2}$. Since current is proportional to potential, the same relationship must apply to the voltage: $V_{max} = V_{eq} \times \sqrt{2}$.

Meters designed for ac are always calibrated to read the equivalent steady dc values I_{eq} or V_{eq}, and loads and ratings are always given in these same terms. Hence your 120-volt ac outlet actually provides a potential difference which rises to a magnitude of $120 \times \sqrt{2} = 170$ volts 120 times/sec.

The sinusoidally varying current I_P in the primary of a transformer causes a sinusoidally varying flux in the core, as shown in Fig. 14-20. At point A the flux is increasing at its maximum rate, and the maximum emf, at a, will be induced in the secondary. At B, the flux is maximum, but at this instant is not changing at all (the flux curve is horizontal), so the induced emf at b is zero. As the flux decreases from B to C to D, a secondary emf is induced in the opposite direction. Calculus could show that the graph of the slope, or rate of change, of a sine curve is also a sine curve, so if we supply a sinusoidally varying input to the primary of a transformer, the output of the secondary will also be sinusoidal.

Each turn of the secondary coil will have induced in it the same emf as every other turn. Since all the turns are, in effect, connected in series, the total emf will be proportional to the number of turns, and the secondary voltage can be made (within certain practical limits) as large or as small as desired by designing the transformer so it has the proper num-

FIG. 14-19 Power in an alternating current circuit.



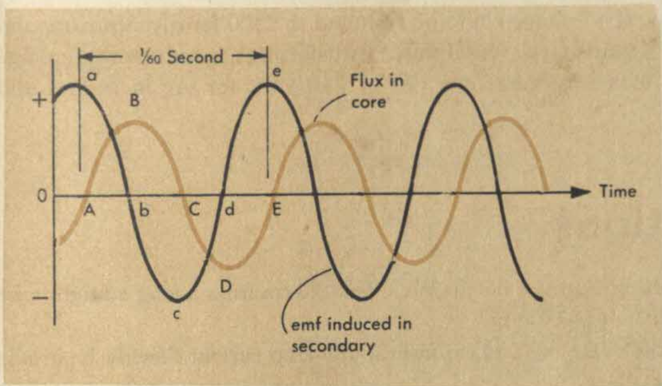


FIG. 14–20 The induction of an alternating secondary emf by an alternating flux in the transformer core.

ber of turns. The relationship between the primary and secondary voltages and the number of turns on the primary and secondary windings is very simple:

$$\frac{V_S}{V_P} = \frac{N_S}{N_P}.$$

One should not assume that the transformer is a slick way of getting something for nothing! Even though the secondary voltage can be made much higher than the input voltage to the primary, the transformer is a remarkable device which automatically regulates the primary current in such a way that the *power* supplied to the primary is always equal to the power taken from the secondary (plus a small fraction more to make up for the inescapable heating losses). In other words, if the secondary voltage is 10 times the primary voltage, the transformer will see to it that the primary current is 10 times the secondary current, so that the input and output powers remain equal.

The ability of the transformer to change voltage is of enormous practical importance. Let us take as an example a generating plant producing electricity at 2300 volts. The plant must supply 4600 kilowatts ("kilo" means 1000, and a kilowatt, or kw, is 1000 watts) to a nearby town over lines which have a total resistance of 5 ohms. To deliver 4,600,000 watts at 2300 volts would require a current of 2000 amperes, and the power used to heat the transmission wires, I^2R, would be 20,000,000 watts! Obviously, this is a very impractical thing to try to do, but engineers can solve the problem by installing transformers to raise the voltage from 2300 to, say, 69,000 volts. To transmit 4.6×10^6 watts at 69,000 volts would require a current of only 67 amperes, and the I^2R losses in the transmission line would be reduced to 22,400 watts. This loss would be less than 0.5 percent of the delivered power, and would be quite economical.

In the city, the voltage could be reduced to 2300 for distribution, and every block would have several small transformers to reduce the voltage still further to a less dangerous 120 or 240 volts for use in homes and offices.

Questions

(14-2)

1. What is the strength of the magnetic field 20 cm from a long straight wire carrying a current of 60 amp?

2. What is the value of H 15 cm from a 200-amp current flowing in a long straight wire?

3. A coil of wire is 8 cm in diameter, and consists of 150 turns; a current of 0.5 amp flows through the coil. What is H at the coil's center?

4. What is the magnetic field strength at the center of a coil of wire consisting of 300 turns, each 12 cm in diameter, if the current in the coil is 0.30 amp?

5. At a point 10 cm from a long straight wire, the magnetic field strength is observed to be 20 amp/m. What current is flowing in the wire?

6. A coil 10 cm in diameter is composed of 50 turns of wire; the magnetic field strength at its center is 120 amp/m. What is the current in the coil?

(14-3)

7. What is the flux density (or magnetic induction) at the point described in Question 1?

8. What is the flux density B at the point described in Question 2?

9. A vertical wire carries a 20-amp current flowing upward. An electron (charge -1.60×10^{-19} coul) 10 cm from the wire is moving upward with a speed of 5×10^6 m/sec. (a) What is the magnitude of the force on the electron? (b) What is the direction of the force?

10. A wire carries a current of 50 amp. An alpha particle moves at 3×10^5 m/sec parallel to the wire and 20 cm from it, in a direction opposite that of the current. What are the magnitude and direction of the force on the particle?

11. A vertical wire carries an upward current. A horizontal beam of electrons is directed toward the wire. In what direction will the beam be deflected?

12. A horizontal E-W wire carries a westward current. A beam of alpha particles is shot directly downward toward the wire. In what direction will the beam be deflected?

13. An alpha particle (mass 6.65×10^{-27} kg) travels at right angles to a magnetic field, with a speed of 6×10^5 m/sec. The flux density of the field is 0.2 weber/m^2. (a) What is the force on the alpha particle? (b) What is its acceleration?

14. An electron (mass 9.1×10^{-31} kg) travels 10^6 m/sec at right angles to a magnetic field whose flux density is 10^{-3} weber/m^2. (a) What is the force on the electron? (b) What is the acceleration of the electron?

15. Since the acceleration of the alpha particle in Question 13 is always perpendicular to its velocity, the speed of the particle will not be changed by it,

although its direction will constantly be changed. (a) Consider the gravitational force on a planet moving in a circular orbit. Do you see a resemblance between the motion of the planet and the motion of the electron? (b) How did we set up the equation from which the radius of the orbit could be calculated? (c) What will be the radius of the path of the alpha particle in the magnetic field of Question 13?

16. An electron (see Questions 9 and 14) is shot into a magnetic field of $B = 2 \times 10^{-3}$ weber/m^2, with a speed of 8×10^5 m/sec. What is the radius of the electron's circular path?

17. An electron is accelerated from the cathode to the anode of an "electron gun." There is a potential difference of 100 volts between cathode and anode, and the electron flies through a hole in the anode into a region containing a magnetic field which is perpendicular to the velocity of the electron. (a) What is the energy of the electron as it passes through the hole in the anode? (b) With this much kinetic energy, what is the speed of the electron as it enters the magnetic field? (c) What must be the flux density of the field in order to make the electron move in a circle of 5-cm radius?

18. An electron gun (see Question 17) accelerates a stream of electrons through a potential difference of 320 volts; the electrons then enter a uniform magnetic field perpendicular to their velocity. (a) With what energy do the electrons enter the magnetic field? (b) With what speed do the electrons enter the magnetic field? (c) What is B, if the electron beam is bent into a circle of 10 cm radius?

(14-4) **19.** A circular loop of wire 4 cm in diameter is placed with its plane perpendicular to a magnetic field of $B = 3 \times 10^{-2}$ weber/m^2. What is the flux through the loop?

20. A circular loop of wire 3 cm in radius is placed with its plane perpendicular to a magnetic field. It is found that there is a flux of 1.5×10^{-5} weber through the loop. What is the flux density of the field?

21. In Question 19, the loop is oriented so that its plane makes an angle of 70° with B. What, now, is the flux through the loop?

22. In Question 20, the loop is oriented so that its plane makes an angle of 65° with B. What must be the magnetic induction of the field if there is to be the same flux of 1.5×10^{-5} weber through the loop?

(14-5) **23.** What are the magnetic field and the flux density within an aircore solenoid 10 cm long, consisting of 1200 turns of wire carrying a current of 0.5 amp?

24. An aircore solenoid 20 cm long and 4 cm in diameter is wound with 5000 turns of wire and carries a current of 0.2 amp. What are the magnetic field and magnetic induction within this solenoid?

25. An iron core (relative permeability = 400 at this intensity of magnetization) is slipped in the solenoid of Question 23. What is the flux density, with the current adjusted to be 0.1 amp?

26. The current in the solenoid of Question 24 is reduced to 0.05 amp, and a core (relative permeability = 750 at this value of H) is slipped in the solenoid. What is the resulting flux density?

(14-6) **27** A wire connecting a dry cell to a lamp is 1 m long and carries a current of 0.5 amp. It is perpendicular to the earth's magnetic field, which has a strength

of 4×10^{-5} weber/m^2. Is there much danger of the wire being ripped loose from its binding posts by the force exerted on the current by the magnetic field?

28. A wire in the armature of an electric motor is 20 cm long, and is in (and perpendicular to) a magnetic field whose flux density is 0.1 weber/m^2. What force is exerted on the wire, which carries a current of 30 amp?

29. A 10-cm length of wire carrying 20 amp is in a magnetic field at right angles to the wire. The wire experiences a force of 0.1 nt. What is the flux density of the field?

30. A wire carrying a current of 50 amp passes through the center of a cylindrical magnetic field 10 cm in diameter. The wire is perpendicular to the field, and experiences a force of 8 nt. What is the magnetic induction of the field?

(14-7)

31. A current of 10^{-7} amp seems quite small. A flow of how many electrons per sec constitutes this current?

32. A certain type of electrometer can measure currents as small as 100 electrons/sec. What is this current, in amperes?

33. A galvanometer with a coil resistance of 100 ohms deflects full scale with a current of 10^{-3} amp. (a) Sketch how a resistance must be connected to make it into a voltmeter. (b) What must this resistance be if the voltmeter is to read 5 volts at full-scale deflection?

34. The coil of a galvanometer has a resistance of 300 ohms and a current of 2×10^{-4} amp causes the meter to deflect full scale. (a) What auxiliary resistance is needed to convert the galvanometer into a voltmeter that reads 25 volts at full-scale deflection? (b) Show how this resistance must be connected.

35. The galvanometer of Question 33 is to be converted into an ammeter reading full-scale for a current of 2 amp. What must be the resistance of the added shunt, and how is it connected?

36. The galvanometer of Question 34 is to be converted into an ammeter reading full-scale for a current of 0.5 amp. What resistance is needed, and how should it be connected?

(14-8)

37. Take two wires, *A* and *B*, separated to 0.1 m. *A* carries a current of 1 amp and *B*, 5 amp in the same direction. (a) Calculate the force produced on *A* by the magnetic field of *B*, and the force produced on *B* by the magnetic field of *A*, per meter of length of wire. (b) How are the directions of these forces related?

38. Two parallel straight wires, *M* and *N*, are 5 cm apart, and each is 0.60 m long. The current in *M* is 30 amp, and in *N*, 5 amp in the opposite direction. (a) Calculate the force on *M* due to the magnetic field of *N*, and the force on *N* due to the magnetic field of *M*. (b) How are the directions of these two forces related?

39. A wire weighing 200 gm/m carries a current of 50 amp, which is fed into the wire by flexible leads whose weight can be neglected. This wire is parallel to, and on top of, a horizontal wire lying on a table. What current must flow through the bottom wire in order to repel the top wire to a height of 3 cm above it?

40. An aluminum wire 50 cm long has a mass of 60 gm, and carries a current of 40 amp, supplied through flexible leads whose weight is negligible. The wire

lies on another wire which is flat on the laboratory bench. What current must flow through the bottom wire in order to repel the top wire to a height of 2 cm above it?

(14-9)

41. A piece of wire is passed through the gap between the poles of a magnet in 0.1 sec. An emf of 4×10^{-3} volt is induced in the wire. What is the flux between the magnet poles?

42. The flux between the poles of a magnet is 5×10^{-4} weber. A piece of wire is drawn across the flux in 0.2 sec. What emf is induced in the wire?

43. A metal airplane with a wingspan of 50 m flies at 360 km/hr. The component of the earth's magnetic field perpendicular to the velocity of the plane is 0.2 gauss. (a) What is the potential difference between the wing tips of the plane? (b) Could you take a lamp, say, and connect it to wires running from each wing tip and have the lamp light up? Explain.

44. A small car whose rear hubcaps are 1.60 m apart is driven 120 km/hr. Where the car is, the vertical component of the earth's magnetic field is 0.3 gauss. (a) What is the potential difference between the hubcaps? (b) Could this potential difference be used to light a small lamp by running wires from the lamp to make contact with each hubcap? Explain.

45. An electromagnet has square pole faces measuring 10 cm $\times$ 10 cm, between which there is a flux density of 3000 gauss. A single conductor of heavy wire is passed across the gap between the poles in 0.5 sec. The ends of the wire are connected to a galvanometer whose resistance is 40 ohms. (a) What average voltage will be induced in the wire? (b) What average current will flow through it?

46. A magnet has pole faces 6 cm in diameter, with a flux density of 5000 gauss between them. A piece of fine wire ($R = 10$ ohms) is connected to a 30-ohm galvanometer, and then passed through the field in 0.2 sec. (a) What average voltage will be induced while the wire passes through the field? (b) What average current will flow through the galvanometer?

47. A small flat coil 2 cm in diameter that has 100 turns and 30 ohms resistance is placed in a magnetic field and then withdrawn. The coil is connected to a ballistic galvanometer, a special galvanometer calibrated to indicate the amount of charge passing through it in a single surge. It has a resistance of 40 ohms, and indicates that a charge 2×10^{-3} coulomb surged through it when the coil was withdrawn. What is the flux density of the field? (The time taken to withdraw the coil makes no difference. Assume it to be Δt; it should cancel out.)

48. A small flat coil of fine wire is 1 cm in diameter, has 50 turns, and a 10-ohm resistance. The coil is connected to a ballistic galvanometer (see Question 47) whose resistance is 50 ohms, is placed in a magnetic field, and is then withdrawn. As the coil is withdrawn, a charge of 3×10^{-4} coulomb passes through the galvanometer. What is the flux density of the field?

(14-11)

49. A 24,000-volt ac transmission line transmits 4.8 megawatts to a town. (a) What is the maximum instantaneous potential difference across the line? (b) What is the effective value of the current in the line? (c) What is the maximum instantaneous current in the line?

50. Some parts of England use 220-volt residential ac electric service. (a) What is the maximum instantaneous potential difference across the contacts

of an electric outlet? (b) What is the effective value of the current in a 4500-watt water heater? (c) What is the maximum instantaneous current in the heater?

51. A transformer has 500 primary turns and 2500 secondary turns. Meters indicate that the load connected to the secondary is operating at 200 volts and 8 amp. What are the voltage and current supplied to the primary?

52. A transformer having 3000 primary turns and 150 secondary turns is supplied from a 240-volt line. The load on the secondary draws a current of 15 amp. What are the secondary voltage and the primary current?

53. A small factory requires 400 kw (kilowatts), and is supplied by lines whose total resistance is 1.2 ohm. How much money would be saved per year by transmitting their power at 11,500 volts instead of 2300? Assume the factory operates 8 hr/day for 300 days/yr, and that electrical energy costs 0.4 cents/kwh. (The kwh, or kilowatt-hour, is an *energy* unit and represents a power of 1000 watts for an hour. A watt-second is, of course, a joule.)

54. A pumping station that operates constantly requires 150 kw of power (see Question 53), which is delivered at 4800 volts over lines whose resistance is 5.2 ohms. How much money would be saved per year by transmitting at 12,000 volts instead of 4800? Electrical energy here costs 0.3 cent/kwh.

chapter / fifteen

Reflection and Refraction of Light

15-1 Reflection of Light

The study of light is one of the most important parts of physics, since most of our knowledge concerning the world around us is gained through seeing. We learn about the properties of giant stellar systems by means of light that travels for millions of years through empty space to deliver us its message. We learn about the properties of atoms through the light emitted by them, which carries in a hidden form information concerning their internal structure. And, of course, most of the information that we get in our everyday life is also obtained through the medium of light.

Unless we are looking directly at a source of light—the sun, a star, an electric light, a firefly, etc.—what we see is reflected light.

Obviously, reflecting surfaces vary in many respects. The shiny door of a well-polished black car may reflect only 5 percent of the light that falls on it, but this small amount of reflected light presents you with a clear image of your face as you stand before it. A good mirror can do no more, although the image is brighter because it may reflect as much as 85 or 90 percent of the light. A piece of black cast iron and a clean

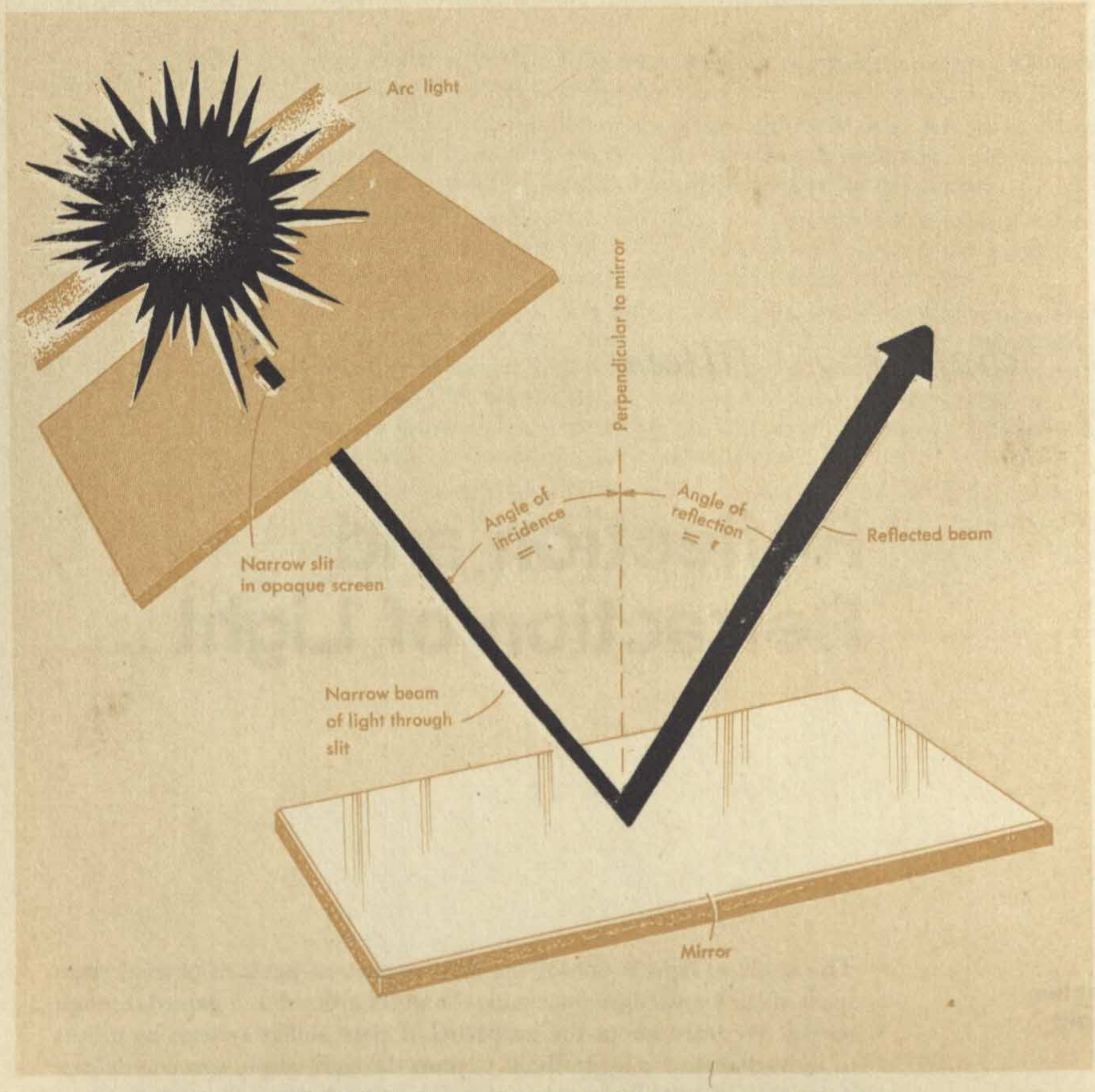


FIG. 15–1 Angle of incidence and angle of reflection for a reflected beam of light.

white plastered wall may reflect as much light as the car door and the mirror, but you see no image of your face as you look at these.

The difference, of course, is in the smoothness of the surfaces. The glossy surfaces of glass or enamel or polished metal are smooth and regular, even to a beam of light; paper and plaster are covered with microscopic irregularities from which light is reflected in all directions like Ping Pong balls bouncing from the surface of a rock pile.

Reflection from a glossy surface is *specular* reflection (from the Latin

word for "mirror"); the irregular reflection from a comparatively rough surface is *diffuse*. Most of our attention will be to specular reflection, because this reflection behaves in accordance with a simple law.

If we direct a narrow beam of light against a mirror (Fig. 15-1) and observe the direction in which the beam is reflected, we can notice a regularity first observed by the ancient Greeks. Draw a perpendicular to the mirror surface where the beam strikes it; the angle between the incoming beam and this perpendicular is called the *angle of incidence*. The angle between the reflected beam and the perpendicular is the *angle of reflection*. The law of reflection is that ***the angle of reflection equals the angle of incidence.*** By applying this single law we can analyze how images are formed by mirrors of various shapes.

15-2 Plane Mirrors

The mirrors that hang on walls are usually flat, or *plane*, and we all are familiar with the appearance of objects when seen in such mirrors. It will probably be helpful to our understanding if the preceding statement is rephrased: instead of "seeing an object in a mirror," it would be clearer to say "seeing the image which the mirror forms of the object." Figure 15-2 shows how the image of an object is formed by a plane mirror. The top of the clover leaf, O (for "object"), is illuminated presumably by the sun; the diffuse reflection of the sunlight from O sends out an infinite number of rays in all directions, of which we have shown only a few. *To the eye, it appears as though the reflected rays were all coming from the point I.* Actually, of course, no light ever reaches or passes through I (I may be in the middle of a thick brick wall), and I is in such a case called a *virtual image* of point O.

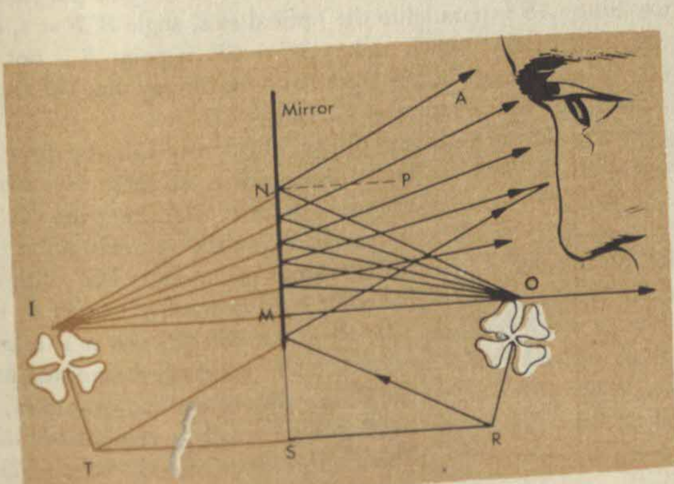


FIG. 15-2 The formation of a virtual image by a plane mirror.

Ray OM has been drawn perpendicular to the mirror, and is reflected back in the direction from which it came. At point N, NP has been drawn perpendicular to the mirror, and for the ray striking N (as for all the rays) the angle of incidence (angle ONP) equals the angle of reflection (angle PNA). The application of a small amount of high school geometry will prove that the triangles OMN and IMN are equal. From this it is apparent that ***in a plane mirror, the virtual image of an object lies on the line drawn from the object perpendicular to the mirror, and is just as far behind the mirror as the object is in front of the mirror.***

The image of the stem of the clover, and of every point between the top and the stem, could have been constructed in this same way, but we need only draw RST perpendicular to the mirror and make TS equal to SR, in order to locate the image of the stem. It is worth noticing that the eye can see the image of the stem perfectly, even though the mirror does not extend down as far as S.

15-3 Concave and Convex Mirrors

In many optical instruments, from telescopes to microscopes, mirrors that are concave or convex rather than plane are used. The simplest shape for such a mirror is that of a small piece of a polished sphere. Let us see what happens to light that is reflected from a concave spherical mirror, shown in Fig. 15-3A. Point C is the center of the hollow sphere of which the mirror is a part and is called the *center of curvature* of the mirror. The line CO, from the center of curvature to the center of the mirror, is the *optical axis*. Take a ray of light which comes in parallel to the optical axis and which strikes the mirror at B. A radius of the sphere CB is perpendicular to the mirror, so AB is reflected toward F with its angle of incidence i equal to its angle of reflection r. The reflected ray hits the optical axis at point F, which is called the *principal focus* of the mirror. Since AB is parallel to the optical axis, angle $BCF = i$, and triangle BCF is isosceles, which makes $BF = CF$. If point B is not too far from O, OF and BF are nearly equal, and we can say that OF almost exactly equals half of the radius of curvature CO.

Many other rays from some very distant object could also be drawn—all parallel to the optical axis and to each other. All these rays would intersect at (or very nearly at) the same point F. The closer the ray AB is to the optical axis CO, the more nearly true is the approximation that $OF = BF$ and that OF equals half the radius of curvature OC. With this restriction in mind, we can define point F as *the principal focus of the mirror; the point at which rays parallel to the optical axis and close to it intersect the optical axis and each other.* The distance OF (generally designated as f) is the *focal length* of the mirror. In all our discussions and examples, we shall assume that the mirror itself will be small in comparison with its radius of curvature, so that these approximations will be very close to the truth.

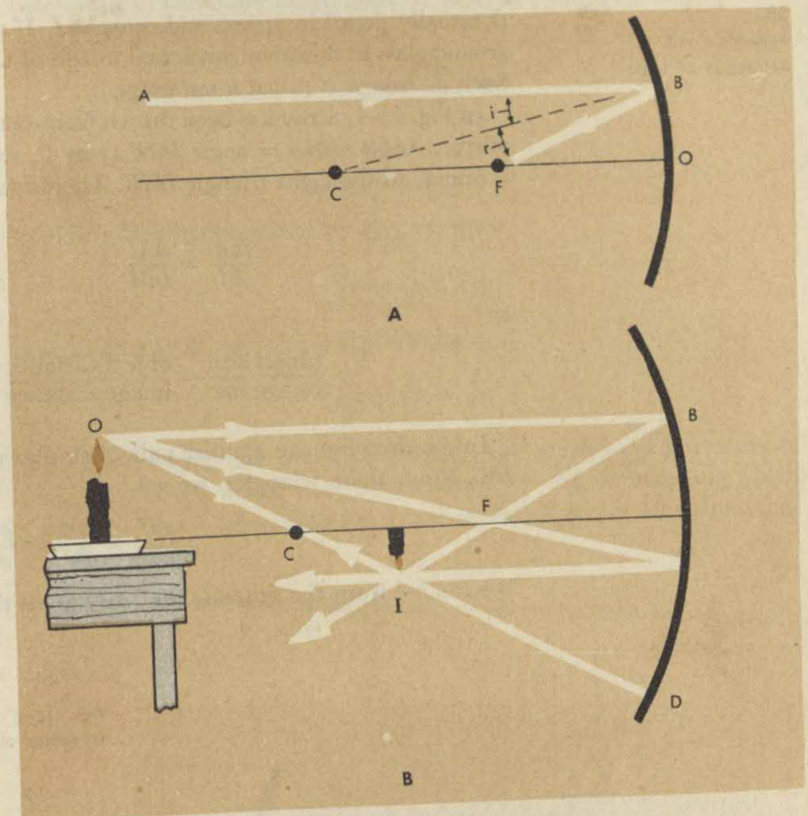


FIG. 15–3 Image formation by a concave mirror.

Figure 15-3B shows the formation of an image by a concave mirror. From point O on the object we can draw an infinite number of light rays which will strike the mirror, and by drawing a radius from C we could construct r equal to i and thus trace the reflection of each. However, by selecting rays judiciously we can make the job much easier. Select OB, the ray parallel to the optical axis; we know this ray will be reflected through the principal focus F. Another ray whose behavior can be easily predicted is OC; since it travels along a radius of the sphere, it will strike the mirror at D and be reflected back toward C again. Since $i = r$, any reflected ray can also be followed backward; in Fig. 15-3A, a ray from F to B would be reflected parallel to the axis along BA. We can thus draw a third predictable ray from point O: the ray OF will strike the mirror and be reflected back parallel to the optical axis, as shown. These three selected rays intersect at I, and to an observer *it will appear as though the light were emanating from I.* The light rays from

O actually do all interesect and cross at I. If we were to put a piece of ground glass at this point, an actual image of O would be focused on it. Such an image is called a *real image*.

In Fig. 15-4, a ray has been drawn from O to M, at the center of the mirror. Angle OMA = angle IMB (why?), so the right triangle OMA is similar to the right triangle IMB. Therefore,

$$\frac{OA}{BI} = \frac{AM}{BM}$$

or

$$\frac{\text{object size}}{\text{image size}} = \frac{\text{object distance}}{\text{image distance}} = \frac{p}{q}.$$

In the drawing are another pair of similar right triangles, OAC and IBC. From these triangles, we get

$$\frac{AC}{BC} = \frac{OA}{BI} = \frac{p}{q}.$$

We can see from the drawing that AC equals the object distance minus

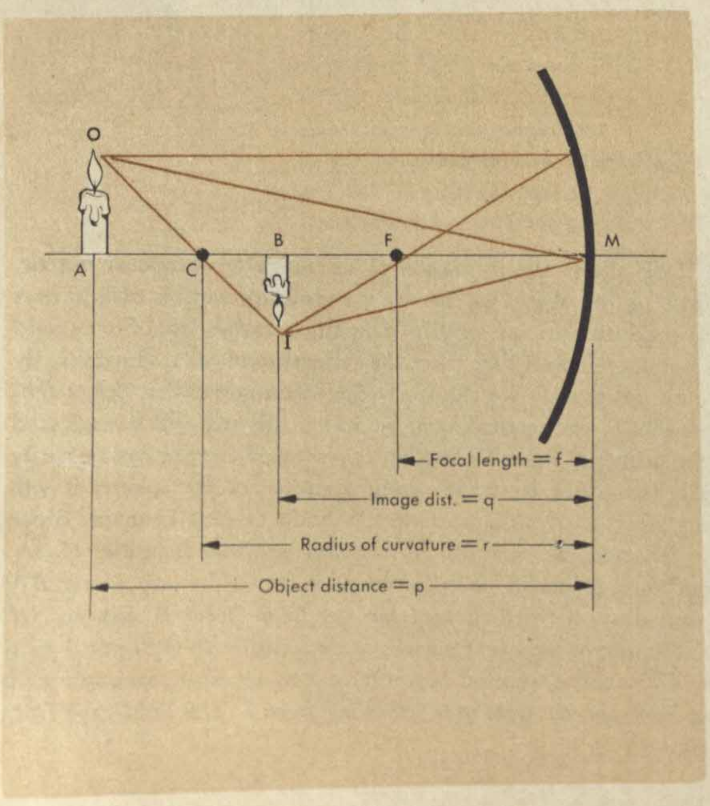


FIG. 15-4 Geometrical factors in the formation of an image by a concave mirror.

the radius of curvature of the mirror, or $p - r$. Similarly $BC = r - q$. Substituting these values we get

$$\frac{p - r}{r - q} = \frac{p}{q}$$

$$pq - rq = rp - pq$$

$$rq + rp = 2pq.$$

Dividing this last equation by pqr, we get

$$\frac{1}{p} + \frac{1}{q} = \frac{2}{r}$$

and since $r = 2f$, and $2/r = 1/f$, we have

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}.$$

As an example, take a concave mirror with a radius of curvature of 24 cm. What can we say about the image of an object 4 cm long which is placed 48 cm in front of the mirror? It is given that $f = 12$ cm and that $p = 48$ cm, so from

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}$$

we get

$$\frac{1}{48} + \frac{1}{q} = \frac{1}{12}$$

or

$$\frac{1}{q} = \frac{1}{12} - \frac{1}{48} = \frac{3}{48} = \frac{1}{16}$$

and

$$q = 16 \text{ cm}.$$

We can now figure the size of the image:

$$\frac{\text{image size}}{4} = \frac{16}{48}$$

from which the image size $= 4 \times 16/48 = \frac{4}{3}$ cm long. The image (as you can check with a graphical construction) will be real and inverted.

It is a good idea to adopt a standard way of making sketches of mirror and lens problems. The most common way is to have the light coming in from the left. If you then keep in mind that the example we have just used to derive the general equation relating p, q, and f is the "all +" case, you should have no difficulty with algebraic signs. The center of curvature of the concave mirror lies to the left of the mirror, and we have called its focal length +; hence, for a convex mirror, the center of curvature will be to its right, and f will be negative. In our "all +" case, the object was to the left; so if the object were to the right of the mirror,

p would be negative. Similarly, for an image formed to the right of the mirror (opposite to the "all +" case), we would have a negative value for q.

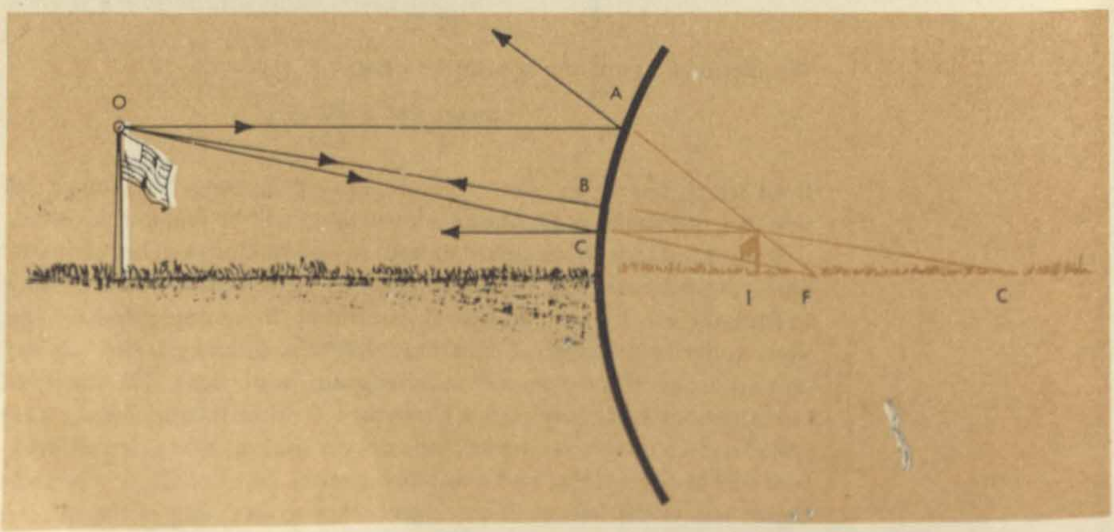
Figure 15-5 shows the graphical construction of the image formed by a convex mirror. The ray OA, parallel to the axis, will be reflected as though it were coming from F. (With a convex mirror, the principal focus F is of course *virtual*, because no ray of light can ever actually reach it.) The ray OB, heading toward the center of curvature along a radius, will be reflected back along the same path. Although the image of the top of the flagpole can be accurately located by the intersection of these two rays alone, we can easily add a third ray: OC is drawn from O toward the principal focus F. It will never get there, of course, but will instead be reflected parallel to the axis and when extended to the right beyond the mirror will also intersect AF and BC at the top of the flagpole image. Thus, to an observer, all three rays (as well as any others we might care to draw) originating at O and reflected by the mirror will appear exactly as though they had come from the virtual image I.

The same equation relating object distance, image distance, and focal length is also applicable to the convex mirror. Suppose an object O were 25 cm in front of a convex mirror with a radius of curvature of -20 cm. The focal length f is then -10 cm, and we have

$$\frac{1}{25} + \frac{1}{q} = \frac{1}{-10} = -\frac{1}{10}$$

$$\frac{1}{q} = -\frac{1}{10} - \frac{1}{25} = -\frac{7}{50}$$

FIG. 15-5 Image formation by a convex mirror.



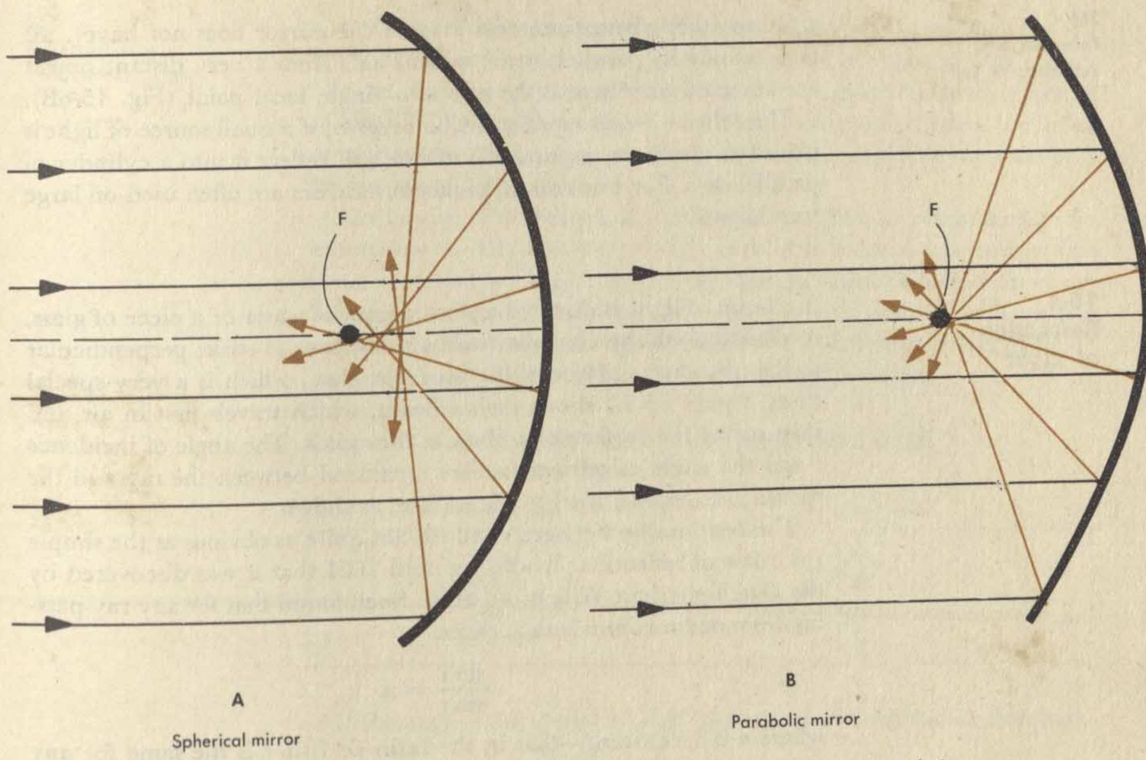


FIG. 15-6 Spherical aberration with a spherical mirror, and its elimination with a parabolic mirror.

or

$$q = -\frac{50}{7} = -7.14 \text{ cm.}$$

The negative value for q shows that the image is behind the mirror and therefore virtual.

Of course, in all the examples considered above, it has been necessary to assume that the actual diameter of the mirror was quite short in comparison with its focal length. Otherwise, the images would be blurred, and we could not even decide exactly where they were located. This is because rays striking the mirror near its rim are reflected to intersect the optical axis closer to the mirror than rays which strike the mirror near its center (Fig. 15-6A). This image imperfection is called *spherical aberration* and is a potential cause of considerable trouble in astronomical reflecting telescopes, whose mirrors must be large in order to collect as much light as possible. This difficulty is solved, however, by making most telescope mirrors parabolic rather than spherical. In such a mirror (which unfortunately is more expensive to make than a spherical mirror

and has other aberrations that a spherical mirror does not have), all rays coming in parallel to the optical axis from a very distant object are reflected to intersect the axis at a single focal point (Fig. 15-6B).

This scheme works equally well in reverse; if a small source of light is placed at the focus, a parabolic mirror will reflect it into a cylinder of parallel rays. For this reason, parabolic mirrors are often used on large searchlights.

15-4 Refraction of Light

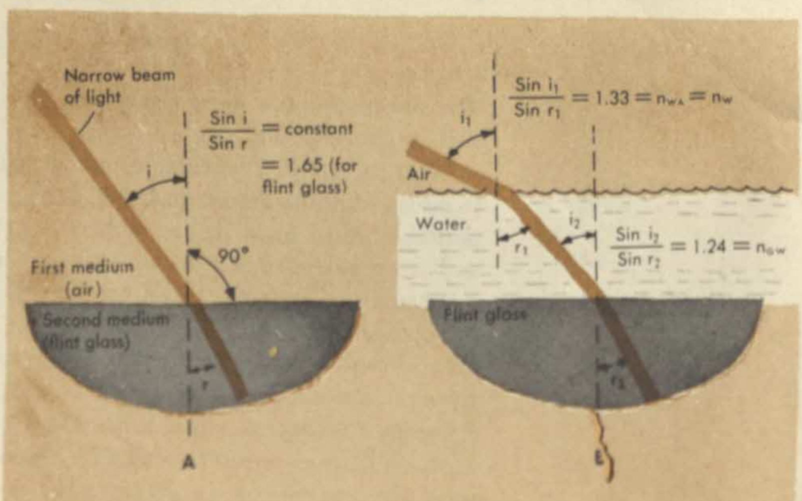
If a beam of light is directed against a pool of water or a piece of glass, its direction will be changed (unless it happens to strike perpendicular to the boundary surface of the water or glass, which is a very special case). Figure 15-7A shows such a beam, which travels first in air and then strikes the surface of a block of flint glass. The angle of incidence i and the angle of refraction r are measured between the ray and the *normal* (i.e., *perpendicular*) to the surface, as shown.

The relationship between i and r is not quite as obvious as the simple $i = r$ law of reflection; it was not until 1621 that it was discovered by the Dutch physicist Willebrord Snell. Snell found that for any ray passing from one medium into another,

$$\frac{\sin i}{\sin r} = n$$

where n is a constant—that is, the ratio $\sin i / \sin r$ is the same for any given pair of media, no matter what the angle of incidence is. This ratio

FIG. 15-7 The refraction of a beam of light in passing from one medium into another.



n is called the *index of refraction* of the second medium (into which the ray is refracted) with respect to the first (from which the ray has come). We can use subscripts to indicate the mediums concerned. In the example of Fig. 15-7A we might write $n_{GA} = 1.65$, to indicate that this is the index of refraction of glass relative to air and would apply to the case of a ray in air entering glass.

Ordinarily, if no reference medium is specified, it is assumed to be empty space. Air, however, differs very little from a vacuum in this respect, so that we might say merely that the index of refraction of the flint glass is 1.65. Table 15-1 gives the index of refraction for a few substances. The index is slightly different for different colors of light; as is customary, these indices are given for yellow light.

TABLE 15-1 REFRACTIVE INDICES OF DIFFERENT SUBSTANCES (for yellow light)

Substance	Refractive Index
Water	1.33
Alcohol	1.36
Glass (flint)	1.65
Diamond	2.42

If the angle of incidence in Fig. 15-7A were 40°, we would then have

$$\frac{\sin 40^\circ}{\sin r} = 1.65$$

$$\sin r = \frac{\sin 40^\circ}{1.65} = \frac{0.643}{1.65} = 0.390$$

$$r = 23.0^\circ.$$

Figure 15-7B shows a situation that is a little more complex; there is a uniform layer of water between the air and the glass. Given the initial angle of incidence i_1, we shall have no trouble in calculating r_1, the angle of refraction into the water, as the index of water (with respect to vacuum or air) is given in Table 15-1. But the table does not give the index of flint glass with respect to water, which we shall need in order to calculate the second refraction. Fortunately (for reasons to be discussed in the next chapter), there is a simple way out of this difficulty: $n_{GW} = n_G/n_W = 1.65/1.33 = 1.24$.

Now, assuming that $i_1 = 60^\circ$, we can work out the example of Fig. 15-7B:

$$\frac{\sin 60^\circ}{\sin r_1} = 1.33$$

$$\sin r_1 = \frac{\sin 60^\circ}{1.33} = \frac{0.866}{1.33} = 0.651$$

and

$$r_1 = 40.6° = i_2.$$

For the water-to-glass refraction,

$$\frac{\sin 40.6°}{\sin r_2} = 1.24$$

$$\sin r_2 = \frac{0.651}{1.24} = 0.525$$

$$r_2 = 31.7°.$$

It is interesting to consider that if the direction of the beam in Fig. 15-7B is reversed, so that the light passes from glass to water to air, the path of the beam will be identical. The i's will become r's, and the r's will become i's; the indices of refraction will now be of water with respect to glass, and of air with respect to water. These will be the inverse of the indices taken the other way and will both be less than 1.

If a transparent solid is submerged in a transparent liquid with the same refractive index, it becomes invisible, since the light rays are not refracted when they pass through the boundary between the two substances. This fact was used by H. G. Wells in his story *The Invisible Man*, which is about a man who managed to make his body transparent by reducing its refractive index to 1. But Wells overlooked, probably intentionally, one essential point: the invisible man would also be blind, since the lenses in his eyeballs would not form any image on the retina.

15-5 Prisms

The prism is a very useful piece of optical equipment, and we shall speak of it again later in connection with optical instruments. Figure 15-8 shows a narrow strip of light falling on a glass prism. If this light is a mixture of red light and blue light, we shall find that the blue is bent more than the red, both on entering the prism and on leaving it. From this observation we must draw the conclusion that the index of refraction of the glass is greater for blue light than it is for red. If we had used a beam of white light to pass through the prism, we would see a continuous spectrum of all colors spread out on the screen, ranging from violet through blue, green, yellow, orange, and red, in that order, with violet bent the most and red the least. This variation of the index of refraction with color is called *dispersion* and is present in all transparent materials. It is responsible for the colored sparkle of diamond rings and dewdrops, and makes it possible for the prism spectroscope to spread a sample of light out so that it can be analyzed. In cameras, telescopes, and microscopes, however, in which we want all colors to be focused together, it is very troublesome.

Let us trace what happens to a narrow beam of white light as it passes through a prism. The prism will be in the shape of an equilateral triangle

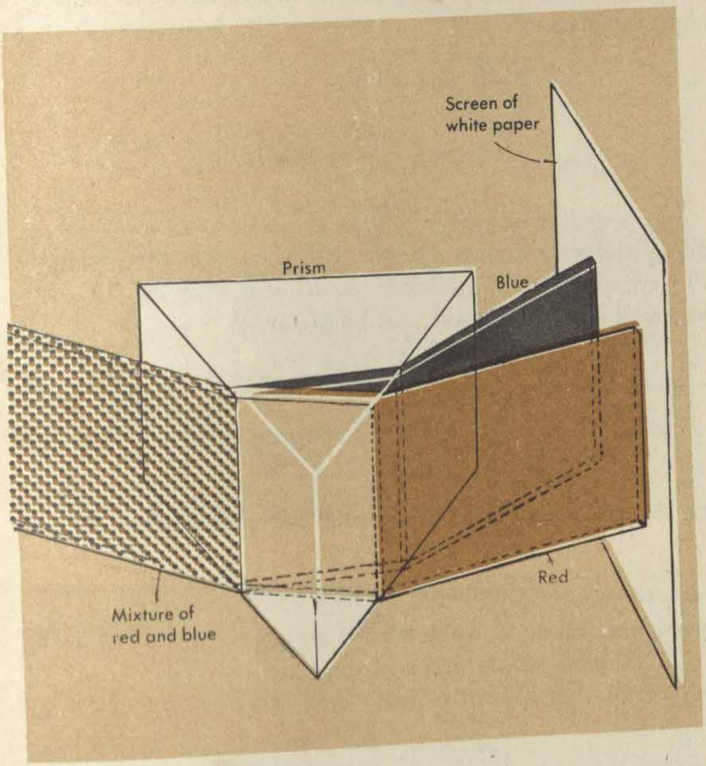


FIG. 15–8 The dispersion of a beam of light as it passes through a prism.

(60°-60°-60°), and will be made of a dense flint glass which has the following indices of refraction for red, yellow, and blue light: $n_R = 1.600$, $n_Y = 1.605$, $n_B = 1.626$. We shall arrange things so that the yellow component of the light passes through the prism parallel to its base BC (Fig. 15-9); in this case, its angle of refraction at the surface AB will be 30°, and the angle of incidence i will be given by

$$\frac{\sin i}{\sin 30^\circ} = \frac{\sin i}{0.500} = 1.605$$

$$\sin i = 0.8025$$

$$i = 53^\circ 22'.$$

Since we have chosen this angle of incidence so that the yellow component of the incident light will pass through the prism perfectly symmetrically with reference to angle A, the yellow component will emerge from side AC into the air with its angle of refraction r' also equal to 53°22′.

The red component of the light will not pass through the prism symmetrically and will require a little more work. Its angle of incidence on

AB will of course be the same 53°22′, and we can calculate its angle of refraction into the glass as

$$\frac{\sin 53°22'}{\sin r} = 1.600$$

$$\sin r = \frac{0.8025}{1.600} = 0.5016$$

$$r = 30°6'.$$

A little adding and subtracting of angles will show that this red component will have an angle of incidence on *AC* of $i' = 29°54'$. Its angle of refraction back into the air will thus be given by

$$\frac{\sin 29°54'}{\sin r'} = \frac{1}{1.600}\left(= n_{AG} = \frac{1}{n_{GA}}\right)$$

$$\sin r' = 0.4985 \times 1.660 = 0.7976$$

$$r' \text{ (red)} = 52°54'.$$

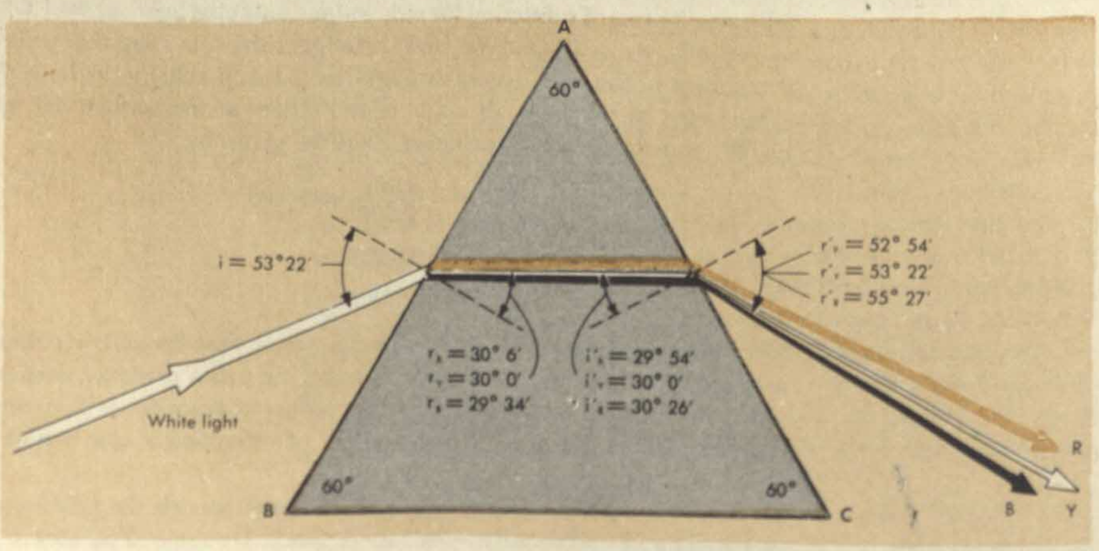
Following the same procedure for the blue component of the ray will show that

$$r' \text{ (blue)} = 55°27'.$$

In this particular example, we thus have an angular spread of 55°27′ − 52°54′ = 2°33′ between the red and the blue.

For every prism that is used to bend and disperse light, there are

FIG. 15-9 Dispersion of a beam of white light passing through a 60° dense flint glass prism.



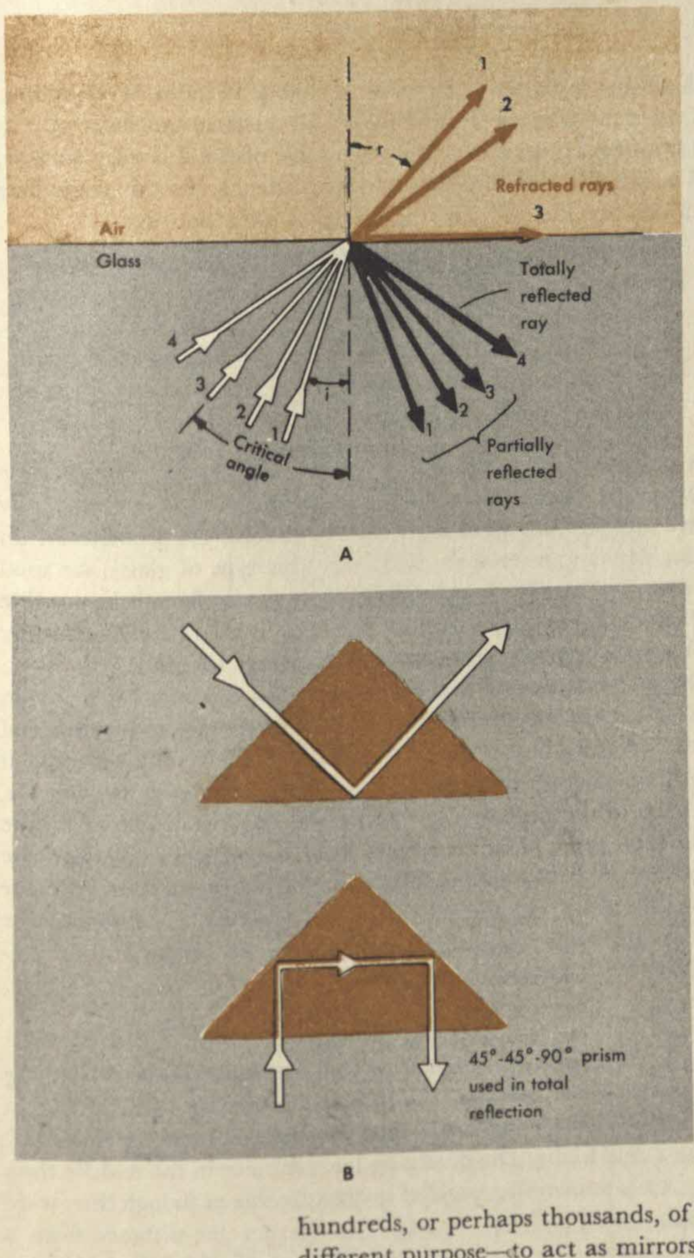


FIG. 15-10 Total internal reflection.

hundreds, or perhaps thousands, of prisms that are used for an entirely different purpose—to act as mirrors in reflecting light. To see how this is accomplished, look at Fig. 15-10A. There are many kinds of optical glass with indices of refraction (for yellow light) varying from 1.45 to 1.92, but let us assume an index of 1.65 for a ray passing from air to glass. The rays in the drawing pass *from glass into air*, however, so the index is $1/1.65 = 0.606$, and we can compute the direction of the refracted rays in air by setting

$$\frac{\sin i}{\sin r} = 0.606.$$

As with all rays when they encounter a change in index of refraction, a part of ray 1 is reflected from the glass-air surface, but most of it is refracted into the air as shown. The behavior of ray 2 is very similar, but ray 3 is quite unusual. If its angle of incidence (for this particular glass) is 37.3°, we can calculate its angle of refraction as

$$\sin r = \frac{\sin 37.3^\circ}{0.606}$$
$$= \frac{0.606}{0.606} = 1$$

and

$$r = 90^\circ.$$

So this particular ray (called the *critical ray*) striking the glass-air surface at this particular angle (called the *critical angle*) cannot really be refracted out into the air at all. If we try to solve the problem for an angle of incidence greater than 37.3° (for this type of glass), we shall find that $\sin r > 1$, which is impossible. Nature also finds it impossible (or *almost* impossible!—as we shall see later), and the ray will be totally reflected back through the glass, as shown in Fig. 15-10A.

Thus the phenomenon of *total internal reflection* can occur if a ray in a material of higher index of refraction tries to escape into a material whose index of refraction is lower. The ray will be totally reflected if its angle of incidence is greater than the critical angle: $\sin i_{\text{crit}} = 1/n$.

In the 45° prisms shown in Fig. 15-10B, the angle of incidence on the glass-air surface is 45°, which is greater than the critical angle, and the rays will be 100 per cent reflected. Besides being a more efficient reflector than any silvered or aluminized surface, there can be no tarnish or corrosion on the glass surface, and it will maintain its 100 percent reflecting ability for an indefinitely long period.

15-6 Lenses

Analogous to concave and convex mirrors are *converging* and *diverging lenses*. These are shown in Fig. 15-11. The converging lens, thicker in the middle than at the edge, bends incoming parallel rays so that they converge to a real focus. The diverging lens, thinner in the middle than at the edge, causes incoming parallel rays to diverge as though they were coming from a virtual focus. As with the mirrors, the distance from a lens to its focal point for parallel rays is called the *focal length* of the lens; for a converging lens the focal length is positive, and for a diverging lens, negative.

With a little juggling about of similar triangles as shown in Fig. 15-11B and following a procedure similar to that which led to a rela-

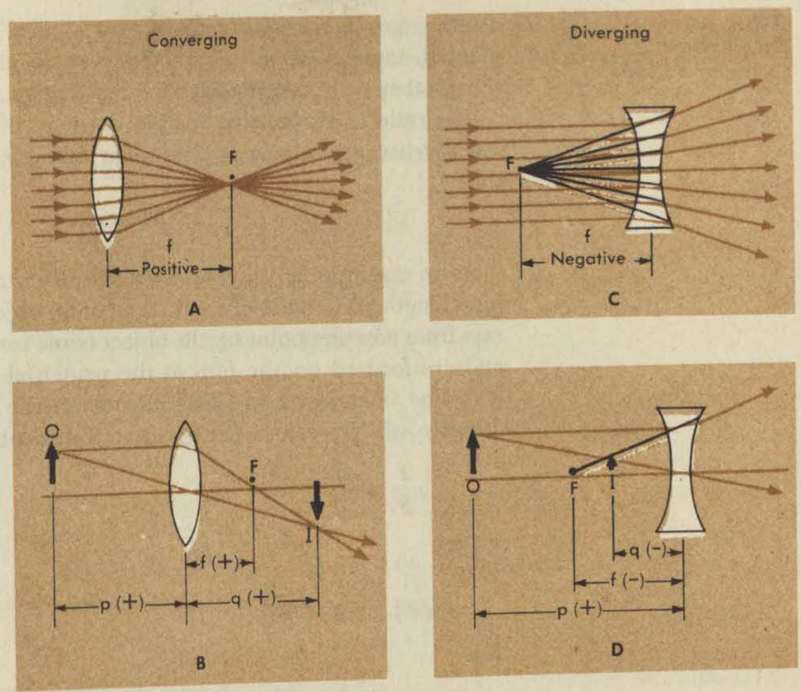


FIG. 15–11 Focal points and image formation by converging and diverging lenses.

tionship between p, q, and f for spherical mirrors, we can derive what turns out to be the identical relationship applicable to lenses:

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}.$$

In order to establish a convention for + and − signs, look carefully at Fig. 15-11B, which is an "all +" diagram. The light proceeds from left to right; the object distance p, to the left of the lens, is positive, and the image distance q, to the right of the lens, is also positive. Since the lens is converging, f also is positive. Figure 15-11D shows the formation of a virtual image by a diverging lens; f is negative, but since the object is to the left of the lens, p is positive. Since the virtual image is to the left of the lens (on the opposite side from the image in our "all +" Fig. 15-11B), q is negative.

The graphical construction of images is as simple for lenses as for mirrors. In Fig. 15-11B, we can pick two of the infinite number of rays we could draw emerging from O. One ray, parallel to the axis, will be refracted to pass through the principal focus. The other we can draw through the center of the lens. At this point the glass surfaces on both sides of the lens are parallel, and since in this elementary work we must

confine ourselves to *thin* lenses, the ray passes through this center point without any deviation. To an observer on the right, the rays look as though they were coming from the real image I, where the rays intersect.

The ratio of image size to object size is, of course, what we call the *magnification*, and a look at the similar triangles in Fig. 15-11B will show that

$$M = \frac{q}{p}.$$

As an example, let us consider a simple camera with a lens of 50 mm focal length. When a picture is taken of an object at a great distance, the rays from any one point on the object come into the lens almost parallel, and are focused on the film at the principal focus 50 mm behind the lens. For an object 1 m (1000 mm) from the lens, however, the image distance will be somewhat greater. Substituting numbers into

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}$$

we get

$$\frac{1}{1000} + \frac{1}{q} = \frac{1}{50}$$

and

$$\frac{1}{q} = \frac{1}{50} - \frac{1}{1000} = 0.020 - 0.001 = 0.019$$

so that

$$q = 52.6 \text{ mm}.$$

Thus, in order that the image of nearby objects may fall exactly on the film, we need to "focus" the camera by increasing the distance between film and lens.

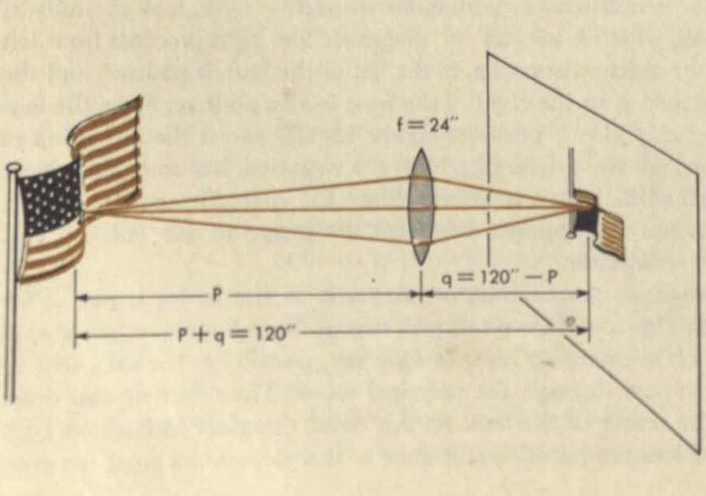


FIG. 15-12 Image formation by a converging lens. The image is reversed left and right and is also upside-down.

Again, suppose we want to focus the image of a flag on a wall 10 ft (120 in.) away from it, and suppose we have a converging lens with a focal length of 24 in. Where must the lens be placed? Figure 15-12 illustrates the problem, and to solve it we must note that $q = 120 - p$ (or $p = 120 - q$; either way will lead to the same solutions). Then

$$\frac{1}{p} + \frac{1}{120 - p} = \frac{1}{24}$$

$$\frac{120}{p(120 - p)} = \frac{1}{24}$$

$$p^2 - 120p + 2880 = 0$$

and

$$p = \frac{120 \pm \sqrt{14,400 - 11,520}}{2}$$

$$= \frac{120 \pm \sqrt{2880}}{2}$$

$$= \frac{120 \pm 53.7}{2}$$

$$= 86.8 \text{ in. or } 33.2 \text{ in.}$$

From this, we see that there are two possible locations for the lens. One will give a magnified image, and the other will produce an image smaller than the flag.

15-7 Lens Combinations

Figure 15-13A indicates a problem which at first glance seems quite complicated—where will this system of three lenses form an image of O, and what will be the total magnification? Like many other problems in physics, a seemingly complicated question turns out to be no more than a series of simple questions. Let us take the lenses one at a time. For lens 1, $1/p_1 + 1/q_1 = 1/f_1$ becomes (Fig. 15-13B)

$$\frac{1}{24} + \frac{1}{q_1} = \frac{1}{6}$$

and

$$q_1 = +8.$$

This image formed by the first lens serves as the object for the second lens, no matter if the image is real or virtual, or whether or not the second lens blocks the rays from ever forming the image at all.

Some adding and subtracting of distances shows that I_1 falls at a distance 4 units to the right of lens 2 (Fig. 15-13C) and becomes a *virtual object* VO_2 for the second lens. Since VO_2 is to the right of lens 2, p is negative, and we have

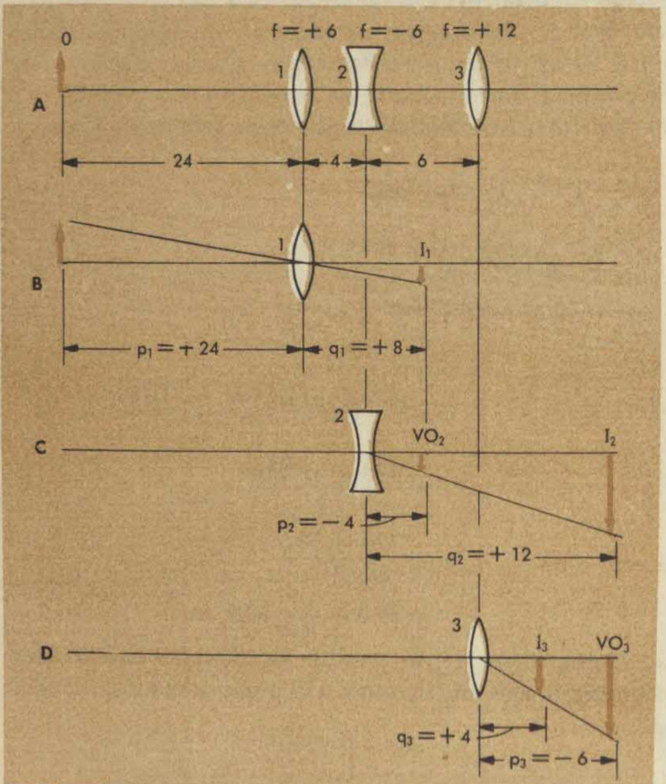


FIG. 15-13 Image formation by a series of lenses.

$$\frac{1}{-4}+\frac{1}{q_2}=\frac{1}{-6}$$

$$q_2=+12.$$

We then make I_2, the image formed by the second lens, the object for the third lens (Fig. 15-13D) and get

$$\frac{1}{-6}+\frac{1}{q_3}=\frac{1}{+12}$$

$$q_3=+4.$$

The final image, then, is formed 4 units to the right of the third lens. To get the total magnification, we need only multiply together the magnifications for each of the separate lenses:

$$M=\frac{q_1}{p_1}\times\frac{q_2}{p_2}\times\frac{q_3}{p_3}$$

$$=\frac{8}{24}\times\frac{12}{4}\times\frac{4}{6}=\frac{2}{3}.$$

Hence the final image will be $\frac{2}{3}$ the length of the original object, and, as is apparent from the drawings, will be real and inverted.

15-8 Microscopes

The most commonly used optical instrument is the simple magnifying glass. As we bring a small object up closer and closer to one eye, it appears larger and larger; if there were no limit to this process, magnifying glasses would be unnecessary. However, there is a limit to the focusing ability of the eye, and for a normal adult, 10 in. or 25 cm is about as close as possible for an object that is to be examined comfortably for any extended period of time. One way of considering the simple magnifier is to think of it as making the eye lens stronger, so that the object may be held closer to the eye and still be focused as a clear image on the light-sensitive retina at the back of the eyeball.

Figure 15-14 shows the use of the simple magnifier, or magnifying glass. If we place the object slightly closer to the magnifier than its focal length, a virtual image is formed at the closest distance for distinct vision, which is assumed to be 10 in. or 25 cm. The magnification is the ratio of the size of this virtual image to the size of the object. To determine the magnification, we have the known f of the magnifier, and we know that $q = -25$ cm. From this,

$$\frac{1}{p} - \frac{1}{25} = \frac{1}{f}$$

$$\frac{1}{p} = \frac{1}{f} + \frac{1}{25} = \frac{25+f}{25f}$$

$$M = \frac{q}{p} = q \times \frac{1}{p} = \frac{25(25+f)}{25f} = \frac{25\text{ cm}}{f\text{ cm}} + 1.$$

If we had taken all our dimensions in inches, it would have come out

$$M = \frac{10\text{ in.}}{f\text{ in.}} + 1.$$

To get greater magnification than is feasible with a simple magnifier, a compound magnifier, or *microscope*, may be used. Figure 15-15 traces three rays from each end of the object through the two lenses of the microscope and into the observer's eye. The first lens, generally of short focal length, is called the *objective*, and forms a real image of the object a short distance in front of the eyepiece lens, or *ocular*. We may consider the eyepiece to be a simple magnifier, through which the image is examined.

15-9 Telescopes

A *telescope* works on the same principle as the microscope. An objective forms a real image of the distant scene, and this image is examined with the eyepiece. In Fig. 15-16, only the undeviated rays passing through the center of each lens have been shown, in order to avoid too many confusing lines. With telescopes, instead of comparing the actual sizes of image

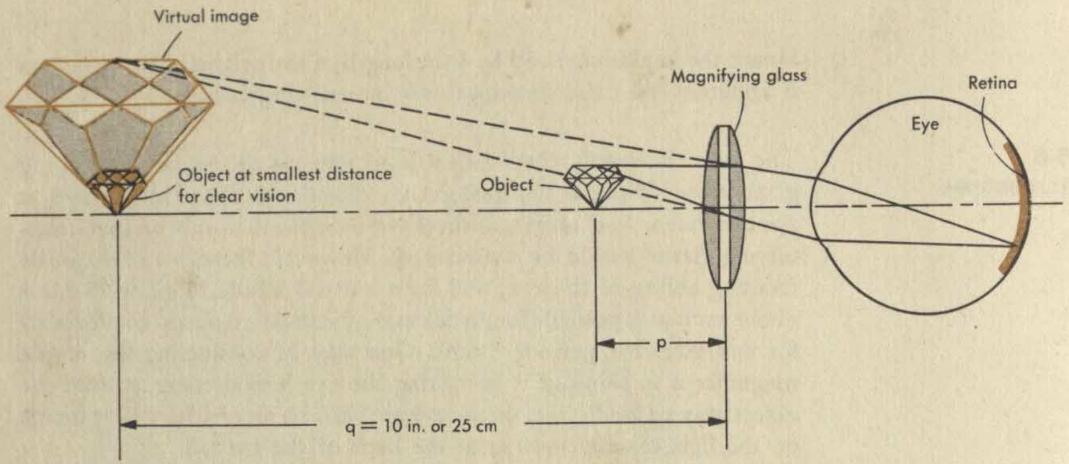


FIG. 15–14 The simple magnifier.

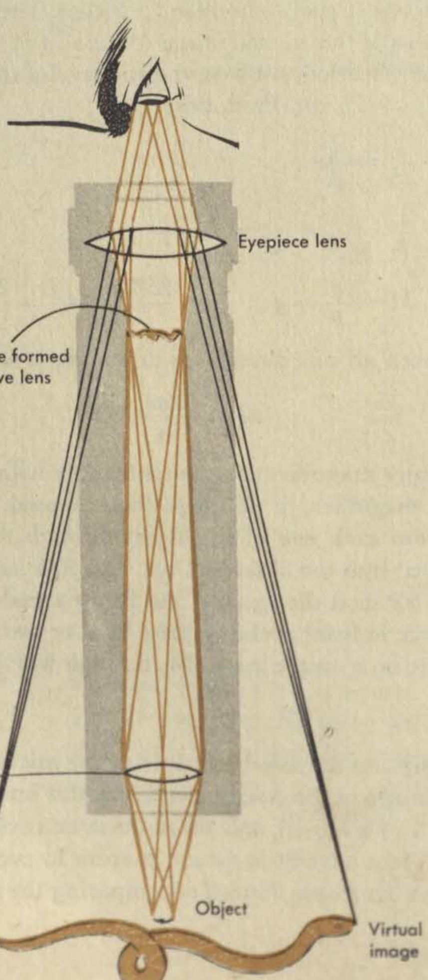


FIG. 15–15 The paths of light rays and the formation of images in a microscope.

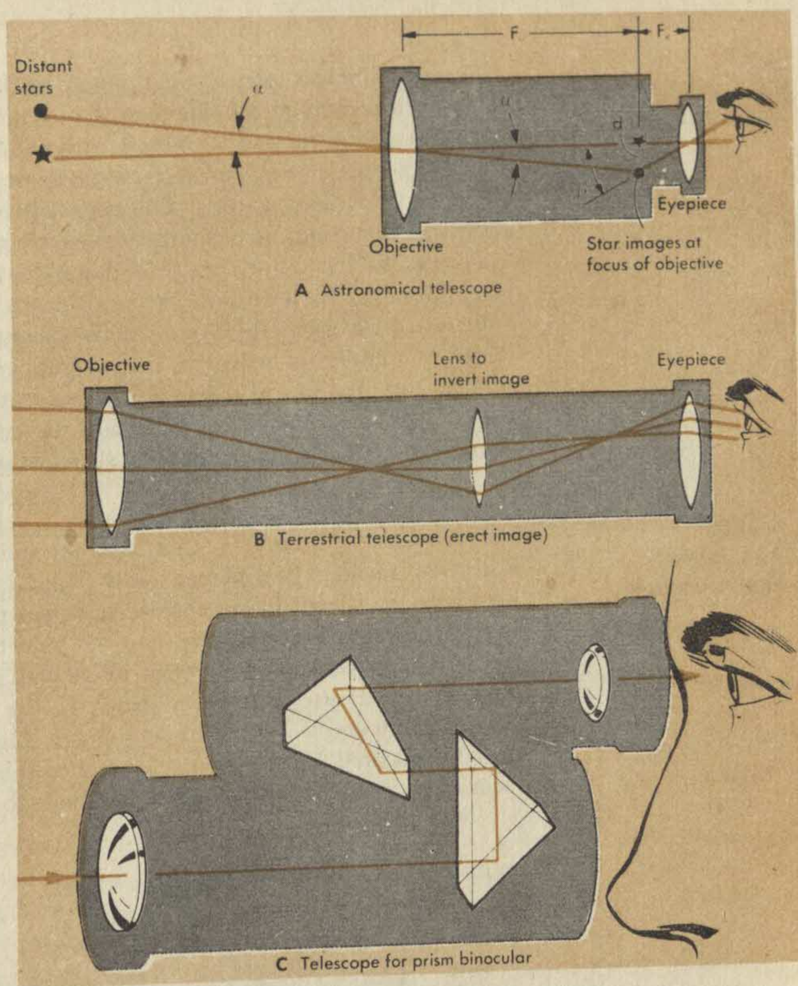


FIG. 15–16 Three types of refracting telescope.

and object, which would be meaningless, we use instead the idea of *angular magnification.* Let us point the astronomical telescope of Fig. 15-16A directly at a star, designated by a star-shaped symbol. Looking up at the sky with no telescope, at an angle α above the star, we can see another star, for which we use a round dot as a symbol. The objective forms real images of these stars, which are separated by a distance d. (If we were to photograph these stars by putting a piece of film at the principal focus of the objective, the star images on the film would be the distance d apart.) Through the eyepiece, the star images appear to be separated by the angle β. To determine the angular magnification, we can express α in radians as very nearly d/F_o while β is d/F_e. The angular magnification, defined as the ratio of β to α, is thus F_o/F_e.

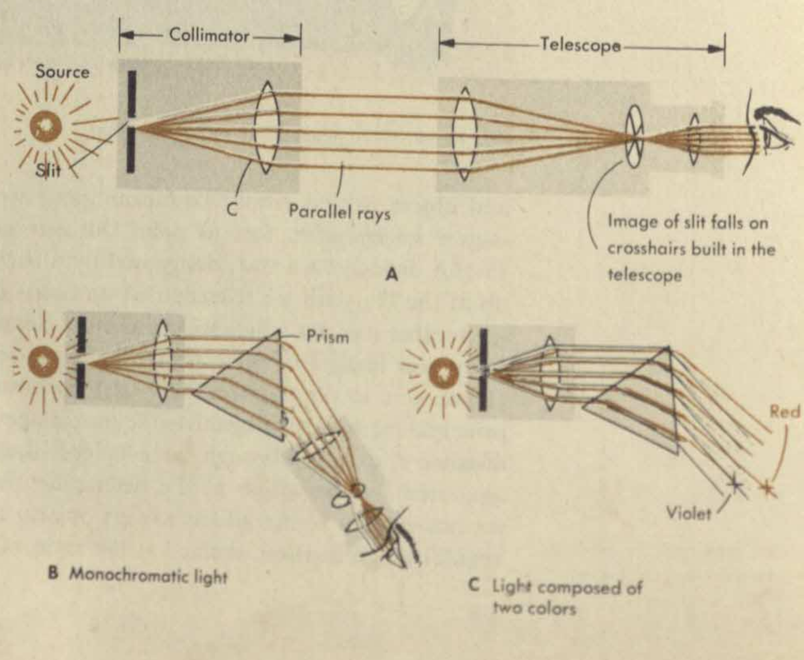
The astronomical telescope has one great drawback which does not bother astronomers at all. Through the eyepiece, the dot-star appears *below* the star-star instead of above it, as it actually is; that is, the image is inverted. This is very inconvenient in a telescope to be used on earthly objects or landscapes, and the difficulty is often corrected by adding an additional inverting, or erecting, lens which puts things right side up again (Fig. 15-16B). The penalty for doing this is that the telescope must be made longer by four times the focal length of the inverting lens.

Binoculars are essentially two telescopes mounted side by side. In order to reduce cumbersome length and weight, *prism binoculars* and *prism telescopes* turn the image right side up and put the left and right sides back where they belong, by reflecting the rays four times from the surfaces of 45° prisms. This scheme is sketched in outline in Fig. 15-16C.

15-10 The Prism Spectroscope

We have already seen (Fig. 15-8) that the index of refraction of glass is different for light of different colors. The fact that different colors of light behave differently has enabled scientists to analyze light, and by their analysis to probe deeply into the structure of atoms and into the nature and composition of stars, as we shall see later. One device for such analysis is the *spectroscope,* shown in Fig. 15-17. Figure 15-17A shows a prism spectroscope without a prism. Light from a source that is to be investigated falls on a narrow slit; the slit is placed at the principal focus of a lens *C* ("*C*" for *collimator,* a device to make light rays or any

FIG. 15–17 The principle of the prism spectroscope.



other things parallel), so that after being refracted by C, the rays from the slit are parallel. These parallel rays fall on a telescope objective which forms an image of the slit at its principal focus. This image of the slit is examined by an eyepiece, as with any telescope. The telescope pivots on a graduated circle, and in order to align it properly, it has crosshairs built in it at the focus of the objective; the telescope is moved until the slit image falls exactly on the crosshairs as seen through the eyepiece.

Now, if we place a prism (usually 60°-60°-60°) between the collimator and the telescope (Fig. 15-17B), the light will be deviated, and we shall have to swing the telescope around to make the slit image fall on the crosshairs again. If the light is all of one single pure color (*monochromatic*), there will be only one slit image. However, if (as in Fig. 15-17C) we use light that is a mixture of two colors, say red and violet, the violet will be deviated more than the red in passing through the prism, and two separate slit images will be formed. To avoid confusion in the diagram, only the crosshairs of the telescope have been shown; in the position marked V, they will be on the violet slit image, and in position R on the red image.

Light such as that which comes from an ordinary incandescent light bulb contains *all* colors, and viewing this light through a spectroscope, we see an infinite number of slit images side by side. These images result in a *continuous spectrum* ranging from the deepest red our eyes can see to extreme violet. In sunlight, certain colors are missing, and if sunlight is allowed to fall on the slit, we see a continuous spectrum crossed by many dark lines, each dark line representing a slit image which is missing because that particular color has been absorbed by the sun's own atmosphere. Such spectra are called *dark-line absorption spectra.* A light such as a neon sign will be seen as a series of bright lines of many colors, most of them red and orange. This indicates that the neon gas emits light of only certain colors; for each precise color, the image of the slit will appear as a bright line of that color, and this spectrum is called a *bright-line spectrum.* In a later chapter, we shall see the reasons why different sources emit these different types of spectra.

Questions

(15-2)

1 A man wants a mirror on a wall, in which he can see his full length. How long must this mirror be? (Consider the man to be a straight line 6 ft tall, with his eye 6 in. down from the top.)

2 In Question 1, just how far up from the floor should the bottom of the mirror be placed? Does it make any difference how far he stands from the wall?

3 Consider two plane mirrors, one on each wall, and butted together in a corner of a rectangular room. Place an object near the corner between the mirrors, and

sketch the images that would be formed, by both single and double reflections. (*Hint:* an image of an image.)

4. A man has a blue left eye and a brown right eye. When he looks in his shaving mirror he sees an image that has a brown left eye and a blue right eye. How would his image appear to him if he looked in the mirrors of Question 3, from a distance great enough that he could not see the single reflections?

(15-3)

5. Take a concave mirror of 8-cm focal length. Compute the image distance when an object is placed at the following distances in front of the mirror: (a) infinite, (b) 48 cm, (c) 16 cm, (d) 12 cm, (e) 8 cm, (f) 4 cm, (g) 2 cm. Which of these images are real, and which virtual? Which erect, and which inverted?

6. The same as Question 5, except that the mirror is convex.

7. A concave mirror has a radius of curvature of 36 in., and a man holds it 12 in. away from his face to examine himself. (a) Where is the image of his nose formed? (b) Is this image real or virtual? (c) If his nose is $1\frac{1}{2}$ in. long, how long is its image? (d) Is the image erect or inverted? Construct the image graphically, as well as making analytical solutions.

8. A concave mirror has a radius of curvature of 24 in. (a) How far from his face should a man hold this mirror if he wishes the image to be formed 12 in. behind the mirror? (b) Will this image be real or virtual? (c) Will it be erect or inverted? (d) If his nose is 3.6 cm long, how long is its image?

9. A polished reflecting sphere 1 ft in diameter is lying on a lawn. A worm crawls toward it. How far from the surface of the ball is the worm when his image appears to be 2 in. behind the ball's surface?

10. A man holds a *convex* mirror whose radius of curvature is 48 cm. How far from his face should he hold it to make his image appear to be 7.2 cm behind the mirror?

(15-4)

11. A ray of light (in air) strikes the surface of a pool of water with an angle of incidence = 45°. What is the angle of refraction of the ray into the water?

12. A light ray makes an angle of 60° with the normal to the surface of a cubical flint glass paperweight. What is its angle of refraction on passing into the glass?

13. A rectangular tank is half-filled with carbon tetrachloride ($n = 1.47$), and the top half filled with water. A ray passing downward through the water strikes the boundary between the liquids at an angle of incidence of 56°. At what angle will it travel in the carbon tetrachloride?

14. In the tank of Question 13, a ray directed upward in the carbon tetrachloride strikes the boundary at an angle of incidence of 32°. What will be its angle of refraction into the water?

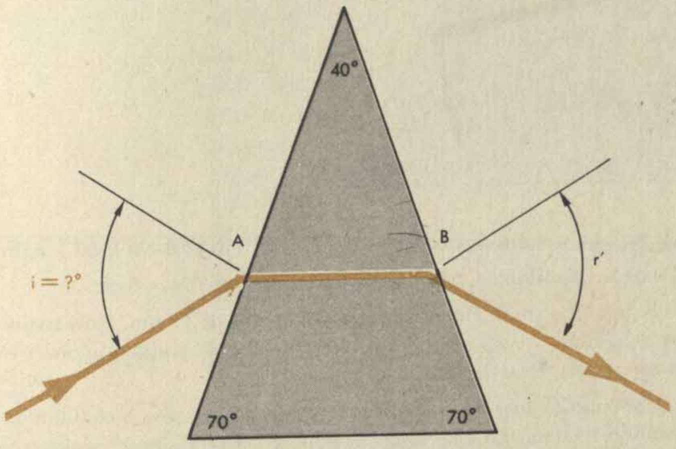
15. A ray of light strikes a thick parallel-sided glass plate with an angle of incidence of 40°. (a) What is its angle of refraction in the glass? (b) What angle does it make with the normal to the surface when it emerges into the air again? (The glass has an index of refraction of 1.60.)

16. A thick parallel-sided slab of glass ($n = 1.60$) is submerged in water. A light ray in the water strikes the slab with an incident angle of 45°. (a) What is its angle of refraction in the glass? (b) What is its angle of refraction from the glass out into the water again?

(15-5) **17.** A ray of monochromatic light strikes a 60°–60°–60° prism with an angle of incidence of 45°. The index of refraction of the prism is 1.500. What is the angle of refraction of the ray into the air after passing through the prism?

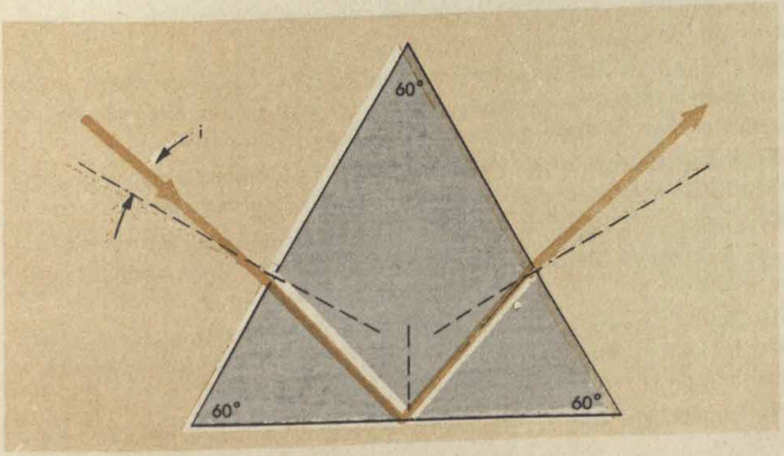
18. The angle between the two refracting faces of a prism is 45°; a monochromatic ray strikes the first face with an angle of 30° *between the ray and the surface.* What is the angle between the ray and the second surface when the ray emerges into the air? ($n = 1.600$.)

19. A ray of yellow light is to pass symmetrically through the prism shown in the figure, that is, the ray AB in the glass is to be parallel to the base of the prism, from which angle r' must equal angle i. What must be the angle of incidence, i, if the index of the glass for the yellow light is 1.600?

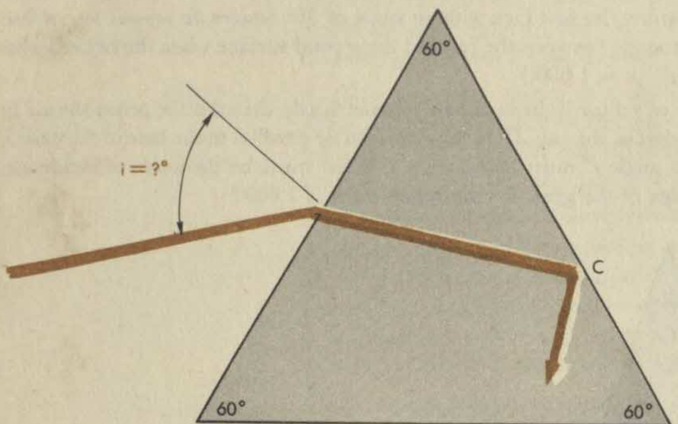


20. Same as Question 19, with a prism whose angles are 75°–75°–30°.

21. A ray is to be totally reflected in passing through a 60°–60°–60° prism, as shown. What is the maximum allowable angle of incidence? ($n = 1.500$.)



22. A 60°–60°–60° prism is made of glass with an index of refraction of 1.414. What must be the angle of incidence of a ray which will be totally reflected at C, as sketched in the drawing?



(15-6)

23. Same as Question 5, but for a converging lens of 8-cm focal length.

24. Same as Question 23, but for a diverging lens of $f = -8$ cm.

25. The lens of a small camera has a focal length of 35 mm. How many millimeters must it be moved out from its "infinity-focus" position in order to focus on an object 2 m from the lens?

26. A "telephoto" lens for a small camera has a focal length of 400 mm. How many millimeters must it be moved out from its "infinity-focus" position to focus on an object 5 m distant?

27. In a movie or slide projector, the brightly illuminated film serves as object, and a lens forms its image on the screen. If the image on the screen is to be 20 times as large as the film in linear dimension (it would have an area 400 times as great) when the screen is 10 ft from the projector lens, what must the focal length of the projector lens be?

28. The lens of a slide projector has a focal length of 5 in. If the image on the screen is to be 30 times as large as the film (a) how far from the film must the lens be, and (b) how far from the lens must the screen be placed?

(15-7)

29. A diverging lens ($f = -10$ cm) is placed 6 cm beyond a converging lens whose focal length is 12 cm. Where will the image of a distant object be formed?

30. Light from a distant object passes through a converging lens ($f = 8$ in.); two inches beyond this lens is a diverging lens ($f = -9$ in.). Where will the image be formed?

31. Parallel light from a distant object falls on a diverging lens ($f = -10$ cm). How far beyond this lens should a converging lens ($f = 25$ cm) be placed, so that the light beyond the converging lens is again parallel?

32. Parallel light from a distant object falls on a converging lens ($f = 6$ in.).

How far beyond this first lens should another converging lens ($f = 9$ in.) be placed so that the light beyond the second lens is again parallel?

33. A converging lens ($f = 8$ in.) is permanently fixed 8 in. from a bright object; 12 in. beyond this lens is a screen upon which a real image of the object is to be formed. You have another converging lens ($f = 6$ in.); where should this second lens be placed in order to form the desired image? (There are two answers to this problem.)

34. A converging lens, $f = 40$ cm, is immovably fixed 40 cm in front of a screen. It is desired to focus on the screen the image of an object 120 cm in front of this lens. You have another lens of $f = 40$ cm available. Where can you place it in order to achieve the desired results? (There are two answers to this problem.)

(15-8) **35.** What is the focal length of a simple magnifier that has a magnification of 10 (often expressed as "10-power" or "10X.") (a) in inches? (b) in cm?

36. What is the focal length (a) in cm, (b) in inches, of a 6X simple magnifier? (See Question 35.)

37. In a compound microscope, the total magnification is the product of the magnification of the objective and that of the eyepiece. The distance from the objective to its real image is fixed by the design of the microscope, and is fairly well standardized at 180 mm. What is the magnification of a microscope using lenses of $f_0 = 4$ mm and $f_e = 25$ mm?

38. What is the magnification of a microscope using objective and eyepiece lenses whose focal lengths are 6 mm and 20 mm, respectively? (See Question 37.)

(15-9) **39.** What is the angular magnification of a telescope whose objective has a focal length of 48 in., with an eyepiece whose focal length is 0.5 in.?

40. An astronomer uses an eyepiece whose focal length is 32 mm on a telescope with an objective focal length of 2.0 m. What is the angular magnification?

41. The "f number" of a telescope or of a camera (generally noted as $f/1.8$, $f/8$, etc.) is the ratio of the focal length of the objective to its diameter. An $f/15$ telescope objective is 6 in. in diameter. Eyepieces of what focal length will be required to provide magnifications of 100, 200, and 300?

42. An $f/8$ telescope (see Question 41) has an objective 4.25 inches in diameter. What magnification will the astronomer have with eyepieces whose focal lengths are $1\frac{1}{2}$ in., 1 in., $\frac{2}{3}$ in., respectively?

43. Excited hydrogen gas may emit a spectrum consisting of bright lines of several different colors, among which are the "C line" (in the red) and the "F line" (blue). A 60° prism is made of glass which has an index of refraction of 1.509 for C, and 1.517 for F. In a spectroscope, light from hydrogen falls on this prism with an angle of incidence of 45°. Through what angle must the telescope swing to go from the C line to the F line?

44. A dense flint glass often used for prisms has $n_C = 1.612$ and $n_F = 1.629$ (see Question 43). Hydrogen light in a spectroscope falls with an angle of incidence = 45° on one face of a 60° prism made of this glass. Through what angle must the telescope swing to go from the C line to the F line?

chapter / sixteen

Light

16-1 The Velocity of Light

In the preceding chapter, we discussed two very important aspects of the behavior of light: reflection and refraction. Nothing, however, has been said about what light *is*. For the moment this need not be a handicap. From the ancient philosophers of Greece and Arabia up through the researches of Newton and into the nineteenth century, scientists of many nations were able to learn a great deal about the behavior of light in spite of their ignorance of its nature.

Light is obviously a form of energy transmitted in some way through space. With your unaided eyes, you can see light that has been traveling from its source for 2 million years to be finally absorbed in the act of stimulating chemical changes in your retinas, which you interpret as "seeing" the Andromeda galaxy.

In order to know that light has taken 2 million years to make this journey, it is apparent that we must have some knowledge of the distance of the Andromeda galaxy (a problem we leave to the astronomers) and also a knowledge of the speed at which light travels.

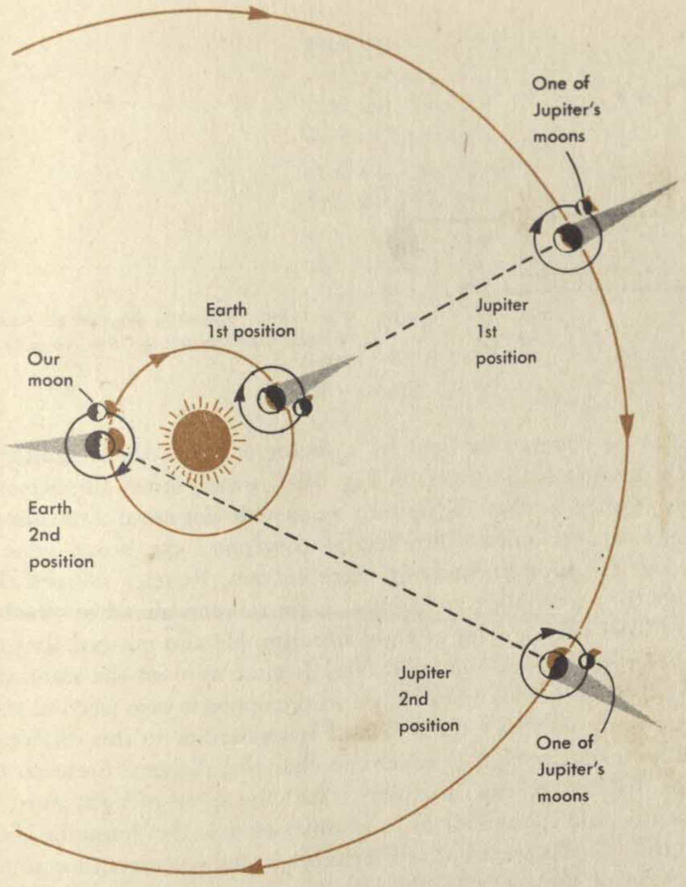


FIG. 16–1 Roemer's method for measuring the velocity of light by observing the eclipses of Jupiter's moons.

The first attempt to measure the speed of propagation of light was undertaken by Galileo in a very primitive way. One evening, he and his assistant placed themselves on two distant hills in the neighborhood of Florence, each of them carrying a lantern with a shutter. Galileo's assistant was instructed to open his lantern as soon as he noticed the flash from the one carried by his master. If light were propagating with a finite speed, the flash from the assistant's lantern would have been observed by Galileo with a certain delay. The result of this experiment was, however, completely negative, and we know now very well why. Light propagates so fast that the expected delay in Galileo's experiment must have been about one hundred-thousandth of a second, which is quite unnoticeable to human senses.

The velocity of light was first successfully measured in 1675 by the Danish astronomer Olaus Roemer, who replaced Galileo's assistant, by observing the moons of the planet Jupiter, thus increasing the

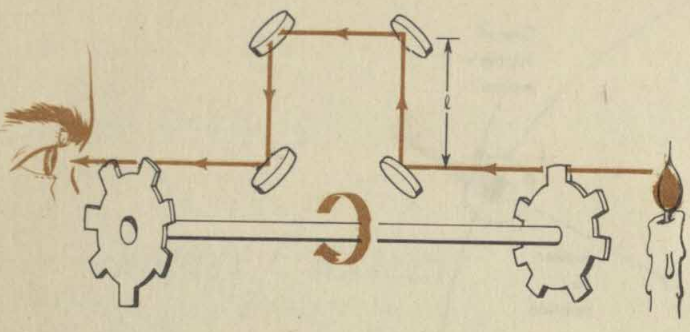


FIG. 16-2 Schematic diagram of Fizeau's method for measuring the velocity of light.

distance to be covered by light by a factor of hundreds of millions.* Roemer's method is illustrated in Fig. 16-1, which shows the orbits of the earth, Jupiter, and one of Jupiter's moons. Moving around the planet, the moons are periodically eclipsed as they enter the broad cone of shadow cast by Jupiter. Studying these eclipses, Roemer noticed that sometimes they took place as much as eight minutes ahead of schedule and sometimes with a delay of eight minutes. He also noticed that the eclipses were early when the earth and Jupiter were on the same side of the sun (first position) and delayed in the opposite case (second position). Ascribing correctly the observed irregularities to the difference in the time taken by light to cover the changing distance between the earth and Jupiter, Roemer calculated that the speed of light must be such that it would require about 16 minutes to cross the diameter of the earth's orbit. Clocks were not very reliable in the seventeenth century, and the size of the earth's orbit was not very accurately known, so Roemer's value of about 2×10^{10} cm/sec was considerably less than the true value as we now know it. His figures did show two things, however: the speed of light was not infinite, as some had believed, but was finite and measureable; and it was inconceivably fast, judged by any earthly standards.

The first laboratory measurement of the speed of light was carried out in 1849 by the French physicist H. L. Fizeau (1819–1896), whose apparatus is shown in principle in Fig. 16-2. It consists essentially of a pair of cogwheels set at the opposite ends of a long axle. The wheels were positioned in such a way that the cogs of one were opposite the intercog openings of the other, so that a light beam from the source could not be seen by the eye no matter what the position of the wheels. However, if the wheels were set in fast rotation and the speed of this rotation was such that the wheels moved by half the distance between the neighboring

* It may be remarked here that Galileo still had a hand in Roemer's measurement of the velocity of light, since it was the former who discovered the moons of Jupiter.

cogs during the interval of time taken by light to propagate from one wheel to the other, the light was expected to pass through without being stopped. In order to observe the effect at the speed of a few thousand revolutions per minute, which was about the maximum that Fizeau could achieve, he had to lengthen the path of the light beam by using four mirrors as shown in Fig. 16-2. This direct laboratory measurement gave a value for the speed of light that was only 5 percent different from the most modern determinations.

If, for example, Fizeau used wheels with 100 teeth spinning at 3000 rev/min we can compute the necessary lengthening of the path of the light rays. While the light traveled to the distant mirrors and back again, the wheels must have turned through the angle between a tooth and a space, or $\frac{1}{200}$ of a revolution. At 3000 rev/min this would require

$$\frac{1}{3000 \times 200} = \frac{1}{600{,}000} \text{ min}, \quad \text{or } 10^{-4} \text{ sec.}$$

We know that the speed of light is almost exactly 3×10^{10} cm/sec, so that in this time interval of 10^{-4} sec, it would have had to travel 3×10^{6} cm, or 30 km. Thus l, the distance to the far mirrors, would have to be about 15 km, or approximately 9 mi. Later, more accurate determinations of the speed of light were made by the American physicist A. A. Michelson, by a similar scheme that used a rotating mirror instead of a toothed cogwheel.

Work is still being done by new and increasingly refined methods, and we can now have full confidence that the speed of light in a vacuum is not far from 2.9979×10^{10} cm/sec. For our work here (in which we generally expect only slide-rule accuracy), we shall simplify our arithmetic by using the value $c = 3 \times 10^{10}$ cm/sec. (The letter c is universally used to represent the speed of light.)

16-2 Photometry

Sources of light can themselves be faint or bright—a firefly or the sun. Illumination can also be faint or intense, from the dim recesses of a dark cave to the bright sunlight of a tropical noon.

Obviously, the intensity of illumination has something to do with the brightness of the source of the light—and equally obviously it also depends on how far away the source of light is. For many years, as was quite natural until this century, the standard of comparison for the brightness of a light source was the *candle*. The standard candle, used by scientists in their measurements, was of a specified size and made of certain fats and oils according to a definite recipe. Even the greatest care, however, could not produce candles that could be depended on to burn with identical brightness.

Modern photometry (the measurement of light and illumination) still uses the words *candle* and *candlepower* (cp), but the standard of refer-

ence is quite different. Roughly, it can be determined as follows: Take a hollow, well-insulated container and have its interior at the temperature of molten platinum just on the point of solidifying (2046°K or 1773°C); now cut a hole whose area is just 1 cm^2. The light emitted through this hole from the incandescent interior is defined to be equivalent to exactly 60 candles. This is a rather difficult apparatus for use by anyone other than the large standardization laboratories, and in actual practice the working standard is generally a carefully made and accurately calibrated ordinary hot-filament incandescent electric lightbulb operated at a specified voltage.

Let us take a 100-watt lamp of about 130 cp. How do we measure the intensity with which it illuminates the page of a book, say, at different distances? We take as the standard unit the *foot-candle*, which is the illumination we would have if the page were held just 1 ft away from a standard candle. So if the book were 1 ft away from the 100-watt lamp, it would be illuminated with an intensity of 130 foot-candles. A problem now might be to calculate the illumination at a distance of, say, 5 ft from the lamp.

Figure 16-3 shows a light source S. At some distance d from the source we place a frame A, 1 foot square; a certain amount of light energy ("*luminous flux*") will pass through the frame, as indicated by the rays marking its boundary. If, now, we place another frame B twice as far from the lamp, its dimensions will have to be 2 ft $\times$ 2 ft in order to just enclose all the luminous flux that has passed through A. So the same amount of light will be distributed over 4 times the area, and the illumination intensity will be $\frac{1}{4}$ as much. Similarly, frame C, at a distance $3d$, must have 9 times as much area as A, and the illumination will be $\frac{1}{9}$ as intense. Accordingly, we see that ***the intensity of illumination varies inversely as the square of the distance from a point source of light.*** (If the source were not small compared with the distances, we could not have drawn such definite boundary rays as those on which the argument was based.)

We can now return to the numerical example which asked for the illumination at a distance of 5 ft from a 130-cp lamp. The solution is now very easy: since the illumination intensity 1 ft away was 130 foot-candles, then 5 ft away it will be only $\frac{1}{25}$ as much, or $130/25 = 5.2$ foot-candles.

The "grease-spot" photometer makes use of the inverse square law of illumination in comparing the brightness of light sources. In its simplest form, this photometer requires as equipment only a piece of white paper with a grease spot in the middle (butter will do very nicely) and a meter stick. A standard lamp of known candlepower is of course also required, as well as the unknown lamp whose candlepower is to be determined. Figure 16-4 shows how the experiment is arranged.

In order to understand how this simple device works, let us first turn

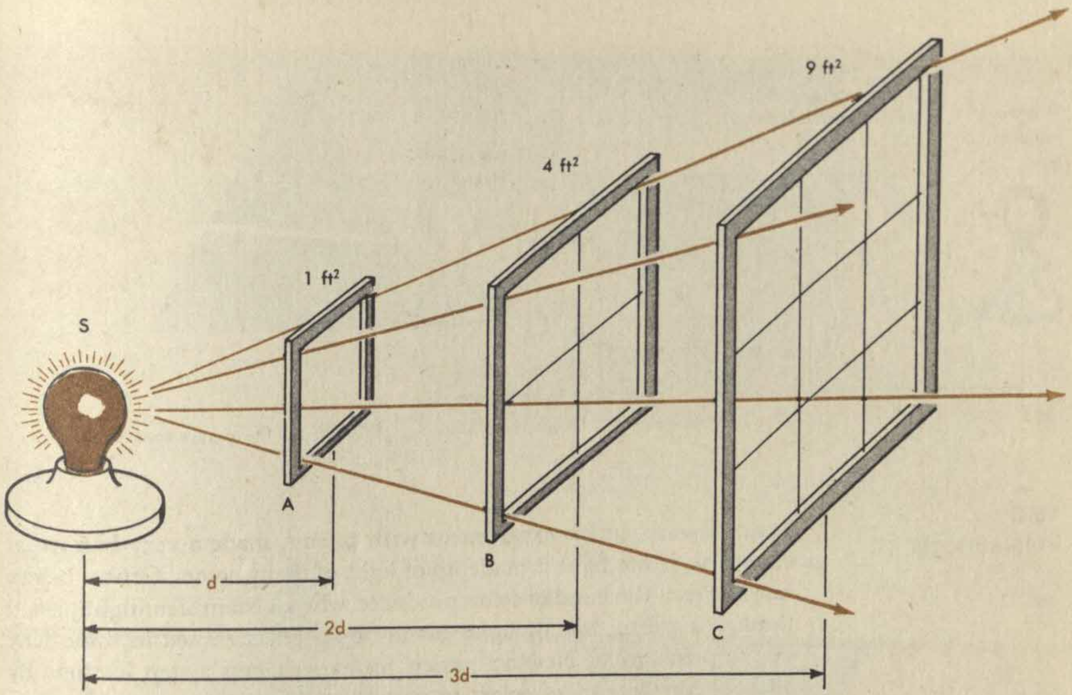


FIG. 16–3 The inverse square law of illumination from a point source of light.

off the standard lamp and view the paper from the right side, illuminated only by the unknown lamp. The paper will appear bright, because most of the light will be reflected from the fibers of the paper. However, where the grease spot is, the paper will appear dark; the spaces between the paper fibers are filled with grease of approximately the same index of refraction as the fibers. Hence there is little or no sharp change in refractive index at the surfaces of the internal fibers, and most of the light will be transmitted through the grease spot, rather than being reflected. This can be confirmed by looking at the left side of the paper; the spot will appear bright from the transmitted light. Now, turn on the standard lamp again and adjust the location of the paper until the grease spot disappears when viewed from either side; the illumination is now equal on both sides of the paper. Using the inverse square law, we can then write

$$\frac{S}{d_s^2} = \frac{X}{d_X^2}; \quad X = \frac{Sd_X^2}{d_s^2}.$$

If, for example, we use a standard lamp of 32 cp and find the spot disappears (or at least nearly disappears, and looks the same from both sides) when d_s is 41 cm and d_X is 59 cm, we have

$$X = \frac{32 \times (59)^2}{(41)^2} = 66 \text{ cp}.$$

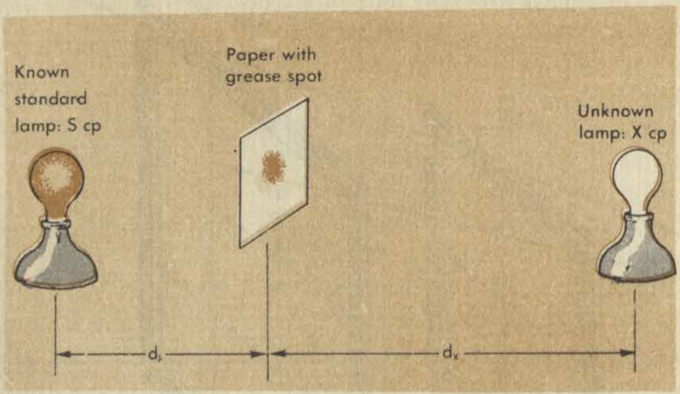


FIG. 16–4 The grease-spot photometer.

16-3 Colored Light

Isaac Newton, in his experiments with prisms, made a very important discovery: white light is made up of light of many colors. Others before him had seen the band of colors produced when a beam of sunlight passed through a prism, but thought this to be an effect caused in some way by the prism itself. Newton carried his experiments a step further. By allowing the spread of colors to pass through another reversed prism (Fig. 16-5A), he found they could be recombined into white again. He also used an opaque screen with a slit in it to isolate various colors of his spectrum (Fig. 16-5B). When he passed monochromatic light of this sort through another prism, he found its color to be unchanged by its passage through the prism. He noted, too, that in this last experiment blue and violet were deviated more than red. He thus not only demonstrated the nature of white light but provided an explanation for the production of a colored spectrum by a prism.

Our lives are filled with color, which is produced in a number of different ways. Let us first consider what we see when we look at a stained-glass window. Outside, the whole window is illuminated by white sunlight; looking from the inside, we see that this light now comes to our eyes in shades of red, green, yellow, blue, etc. Has the glass added some quality of "color" to the light passing through it? With Newton's experiments in mind, we know this is not the case; the stained glass merely absorbs some of the colors from the white light falling on it, and we see what is left.

Figure 16-6 shows the results of passing white light through some hypothetical filters of colored glass. The resulting colored light has supposedly been examined (or better, photographed) through a spectroscope, with the results shown. Filter *a* has absorbed the red, orange, yellow, and green components, leaving only the blue and violet to be transmitted through. (Our judgment of colors is quite subjective, and we cannot place definite

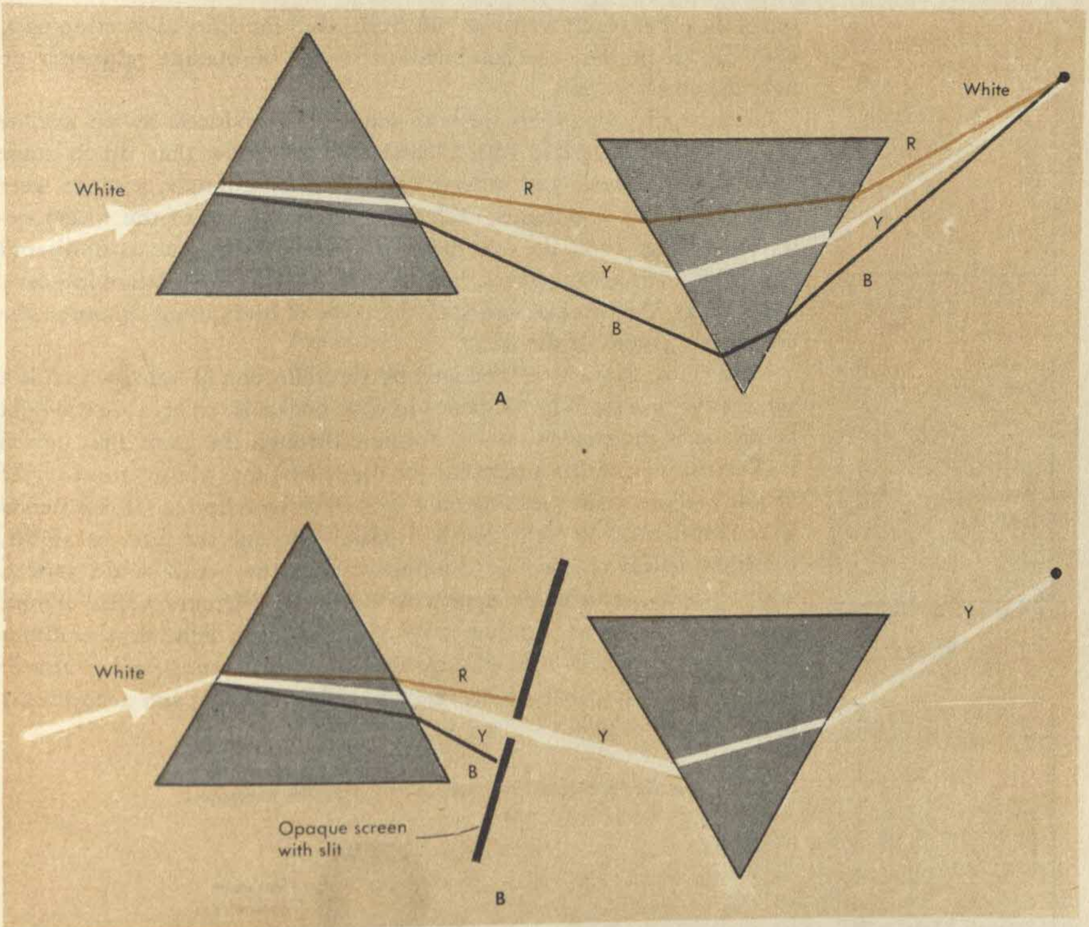


FIG. 16-5 Recombination of spectral colors into white light.

boundaries to separate what we would call green from what we would call blue. In the next chapter, we shall be able to describe this aspect of light more quantitatively; for now it will be enough to merely name colors in the order in which they are spread out by a prism spectroscope.) Similarly, filter *b* absorbs red, orange, blue, and violet, allowing only yellow-green to pass through. And the color transmitted by filter *c* is an orange-red.

The production of colors by these filters is a *subtractive* process. Certain color components have been subtracted by absorption from the white light that fell on them originally. The filters add nothing; they can only take away. If we were to put any two of the filters together, one on

top of the other, each would absorb from what the other allowed to pass, with the result that the combination would be opaque, allowing no light to come through.

Occasionally, however, we may see colors produced by an *additive* process. Those familiar with theatrical work know that much stage illumination is done with colored lights. If, for example, a scene were lighted by three floodlights, each one covered by one of the filters described in Fig. 16-6, the overall result would be the same as unfiltered light. With this arrangement, however, by varying the relative intensity of his lights, the director can vary the color of his lighting to match the emotional "color" of the scene.

Most color, though, is produced by the reflection of light, which is a subtractive process. The red pencil in your pocket is red because the light falling on it penetrates a short distance through the paint that covers it. The paint contains a pigment (or pigments) that absorb most of the yellow-to-violet color, leaving only the red to be reflected. If this pencil were illuminated by light that had passed through the filter in 16-6A, the light would contain no component that the pencil could reflect, and it would accordingly appear to be black. Mercury vapor lamps, often used for street lighting, give a bluish-green light that contains almost no red. Under this lighting, most shades of lipstick (which absorb blue and green and reflect only red) can reflect nothing, and accordingly appear nearly black.

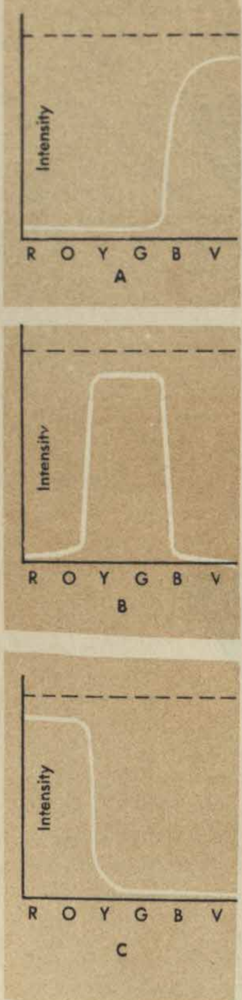
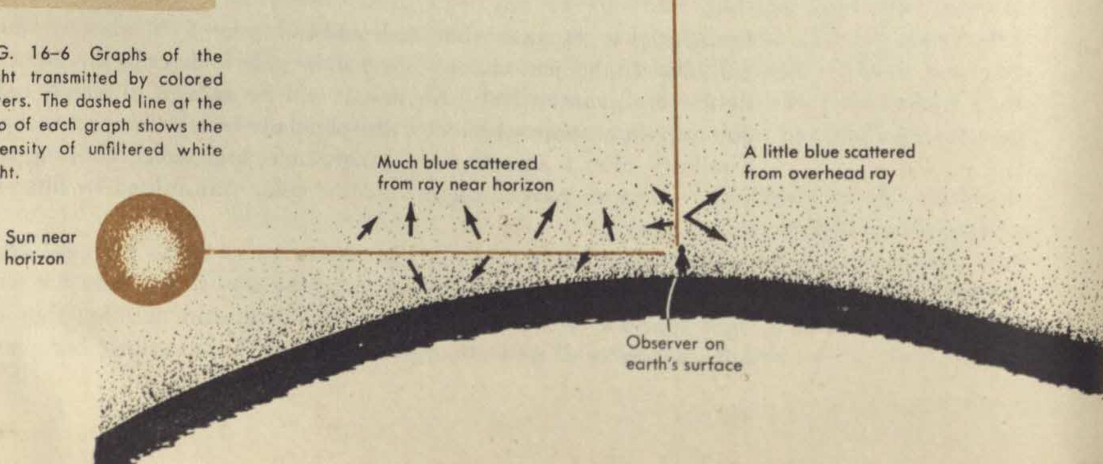


FIG. 16-6 Graphs of the light transmitted by colored filters. The dashed line at the top of each graph shows the intensity of unfiltered white light.

FIG. 16-7 Scattering of blue from the sun's rays by the atmosphere.



16-4 Color in the Sky

From earliest childhood, we all are so accustomed to the blue color of clear sky that it seldom occurs to us to wonder why this should be so. We see the sun, and it is easy to visualize the rays coming from the sun into our eyes. We see the shadows of trees and buildings, and it is also easy to visualize the sun's rays that mark the shadow edges. But in the sky, where there are no such markers to make them visible, it is all too easy to forget that the direct rays of the sun are also passing through every cubic millimeter of the atmosphere, wherever we choose to look.

The sky looks bluest when it is cleaned of dust and smoke, as it often is after a good rainstorm. It is also then very transparent—but *not quite perfectly transparent!* The molecules of the air (which we shall say more about a few chapters further on) present tiny obstacles to the free passage of the light. We can picture some of the light bouncing off these molecular obstacles in all directions—in other words, some of the incoming light from the sun is *scattered* by the air molecules. Now for reasons that are too mathematical for us to go into here, light of short wavelength is scattered much more than light of long wavelength, so that the blue end of the spectrum is more susceptible to this scattering than the red end. Thus, wherever we look in the sky, we see the blue light that has been scattered out of the white sunlight passing through it.

Big particles such as dust and the water droplets that form clouds—enormously larger than the molecules of the air—have very little of this selective effect, and scatter or reflect all colors almost equally. Thus clouds are white, and when the atmosphere is dusty the blue may be nearly obscured by the overall white skylight.

This selective scattering of the blue end of the spectrum has an effect on the light that comes directly to our eyes. Even at noon, when the sun is most nearly overhead, it appears to be yellowish, rather than white. This, of course, is because some of the blue has been scattered out during the passage of the light through the atmosphere over our heads. At sunset or sunrise (or moonset or moonrise) this effect can often be seen in exaggerated form. When the sun is near the horizon, its rays must pass through much more atmosphere to reach us (Fig. 16-7). More of the blue is scattered out of the direct beam; if atmospheric conditions are right, the setting sun may appear to be almost red.

16-5 The Rainbow

When a rainstorm passes and the sun shines again in the sky, we often see a brightly colored arch against the dark background of the departing clouds opposite to the sun. The rainbow is an optical phenomenon caused by the reflection and refraction of sunlight in the tiny spherical raindrops on which it falls. In order for the direction of the rays of sunlight to be nearly reversed, the rays must enter the raindrops and be reflected from their inner surfaces as indicated on a greatly exaggerated scale in Fig. 16-8. The ray marked W indicates a thin beam of white light coming

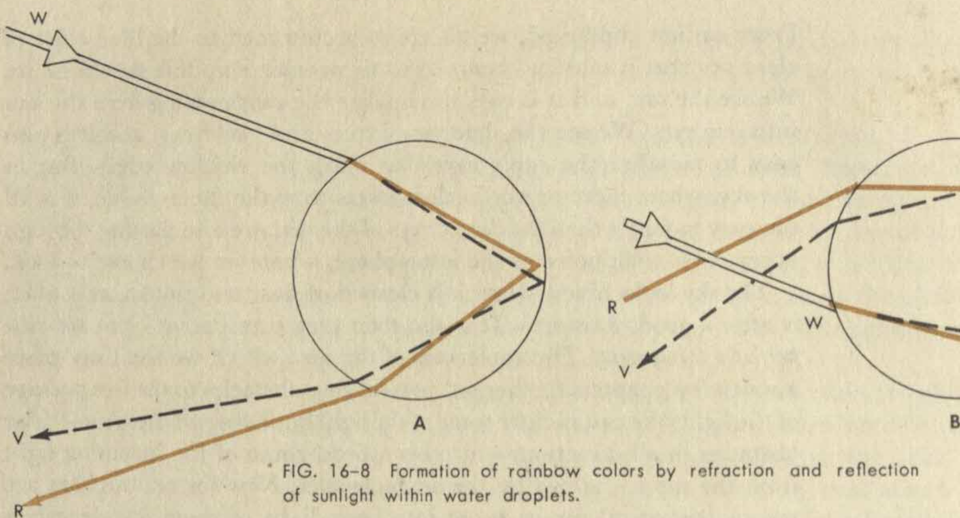
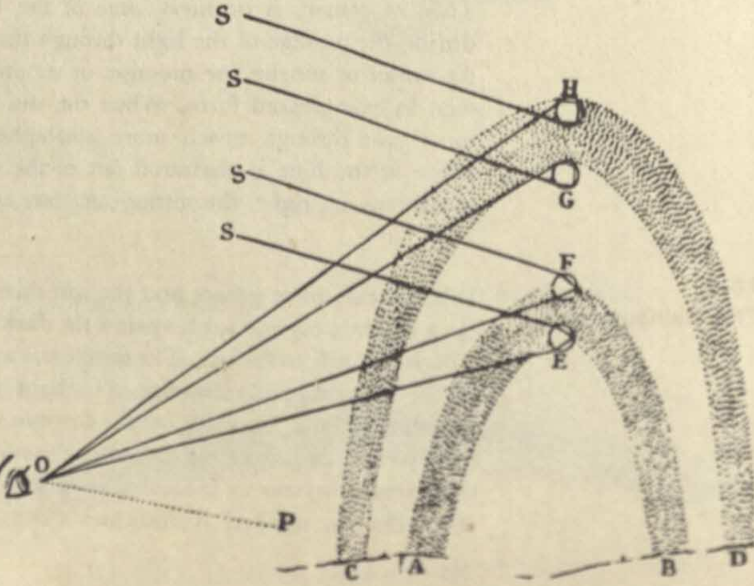


FIG. 16-8 Formation of rainbow colors by refraction and reflection of sunlight within water droplets.

FIG. 16-9 Newton's own explanation of a rainbow: "Suppose now that O is the spectator's eye, and OP a line drawn parallel to the sun's rays and let POE, POF, POG, and POH be angles of 40 degr. 17 min., 42 degr. 2 min., 50 degr. 57 min., and 54 degr. 7 min., respectively, and these angles turned about their common side OP, shall with their other sides OE, OF, OG, and OH describe the verges of two rainbows AF, BE, and CHDG. For if E, F, G, and H be drops placed any where in the conical superficies described by OE, OF, OG, and OH, and be illuminated by the sun's rays SE, SF, SG, and SH; the angle SEO being equal to the angle POE, or 40 degr. 17 min., shall be the greatest angle in which the most refrangible rays can after one reflection be refracted to the eye, and therefore all the drops in the line OE shall send the most refrangible rays most copiously to the eye, and thereby strike the senses with the deepest violet color in that region." From Newton's Optics. (Reprinted by G. Bell & Sons Ltd., London, 1931.)



from the sun. The solid line in Fig. 16-8A shows how the red component is refracted as it enters the drop, is then reflected from the back of the drop, and is refracted again as it leaves the drop to return to the air. This red ray as it leaves makes an angle of very nearly 42° with the direction of the entering white ray.

We have just described the ray *SFO*, as shown in Newton's own diagram, Fig. 16-9. The line *OP* is drawn from the eye directly away from the direction of the sun; if we look 42° away from this line in any direction we shall see the circular arc of red in the sky wherever there is a curtain of water droplets to reflect and refract this color back to us. Returning to Fig. 16-8A, we see that the violet component (Newton's "most refrangible rays"), emerges from the drop, making an angle of only 40° with the direction of the sunlight. Hence if we describe a circle in the sky 40° in radius as measured from the point opposite the sun, it will mark the smaller violet arc, lying inside the red one.

Figure 16-8B shows the double reflection within the droplets, which causes the larger secondary bow we can sometimes see.

16-6 Why Is Light Refracted?

Why do light rays change their direction when they travel from air into water or glass? This question has an important bearing on the problem of the nature of light. Sir Isaac Newton believed that a light beam represents a swarm of tiny particles that are emitted from light sources and fly at high speed through space. He visualized the refraction of light rays entering any material medium as being caused by a certain attractive force acting on the particles of light when they cross the surface of a material body (Fig. 16-10A). The vector diagram illustrates Newton's idea of refraction. In it, v_A represents the velocity of the light in air, and Δv is the increase in velocity due to the attractive force at the boundary. The resultant velocity of the particles in glass is given by v_G, which differs from v_A in both magnitude and direction. It is important to realize that, according to Newton's views, the velocity of light in substances with a high refractive index should be larger than the velocity of light in air or in a vacuum, because the force pulling the light particles into the denser materials adds to their velocity.

Newton's views were opposed by the Dutch physicist Christian Huygens (1629–1695), who believed light to be a wave motion in a certain all-penetrating medium ("world ether") and the propagation of light to be similar to the propagation of sound waves through the air. It is interesting to notice that the explanation of the refraction of light based on Huygens' ideas leads to conclusions concerning light velocity in dense media that are directly opposed to those reached by Newton. To understand it in a simple way, let us substitute, for successive waves of light approaching the surface of glass, successive waves of tanks operat-

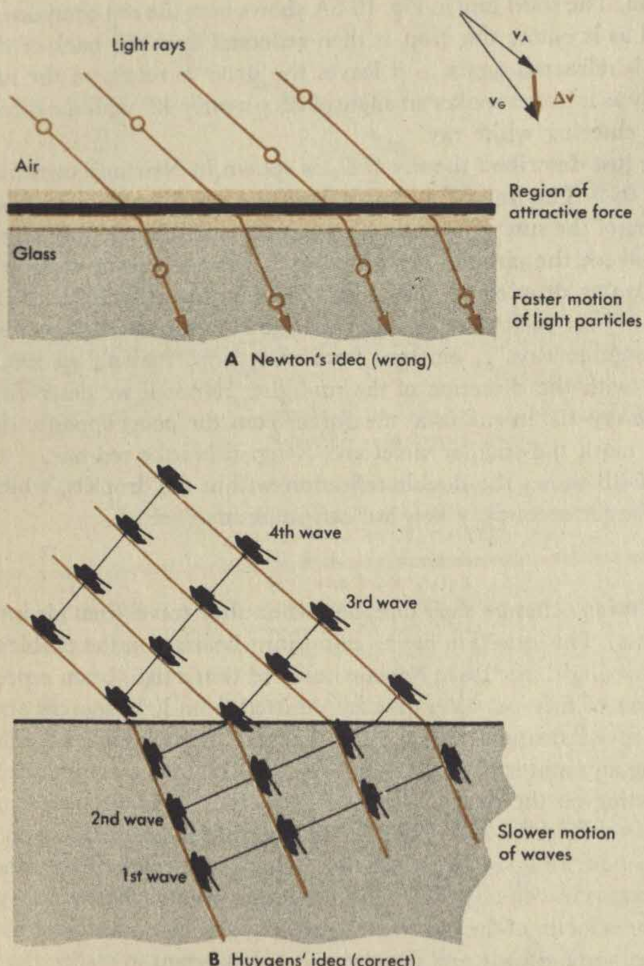


FIG. 16-10 Newton's and Huygens' explanations of the refraction of light.

ing in open country. As the tanks cross a boundary between comparatively good terrain and poor, sandy terrain, their velocity is reduced (Fig. 16-10B). If, after getting into the sandy terrain, individual tank commanders stubbornly maintain their original orderly alignment, the advancing line of tanks that entered the sandy terrain earlier will be delayed in their advance with respect to those that enter it later. To make the rows of tanks continue to be at right angles to their direction of advance, it becomes necessary to change the course of advance and turn all the tanks somewhat to the right.

We can analyze this quantitatively by using Huygens' principle that each point on a wave front could itself be considered a source of further

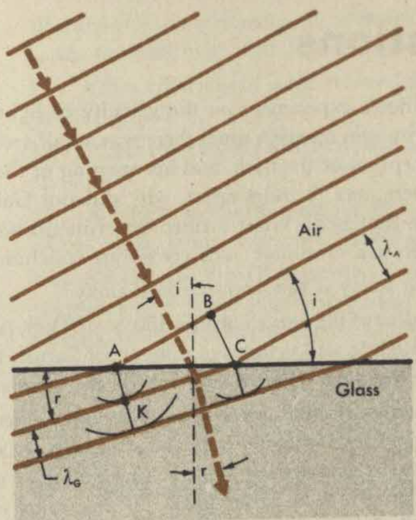


FIG. 16–11 Wave fronts entering glass from air.

waves. Figure 16-11 shows a series of wave fronts traveling in the direction of the heavy dashed line. At point A, one of the wave fronts is just entering the glass, and another point B on the same wave front is only one wavelength away. During the time that B travels to C, the disturbance at A will have traveled in the glass a shorter distance AK, and the wave front in the glass will be represented by KC. Since BC and AK were covered in the same length of time, the velocities in the two media are in the same ratio as these distances, or

$$\frac{v_A}{v_G} = \frac{BC}{AK}.$$

From the drawing, it is plain that $AK = AC \sin r$ and that $BC = AC \sin i$, so we have

$$\frac{v_A}{v_G} = \frac{AC \sin i}{AC \sin r} = \frac{\sin i}{\sin r} = n_{GA}.$$

We thus see that, according to the wave theory, the index of refraction actually represents the ratio of the velocities of light in the two media concerned. We know, for example, that the speed of light in air or vacuum ($n = 1.00$) is 3×10^{10} cm/sec; therefore, in diamond ($n = 2.42$), the speed of light is $3 \times 10^{10}/2.42 = 1.24 \times 10^{10}$ cm/sec.

In 1850, the French physicist Jean Foucault used a rotating-mirror modification of Fizeau's apparatus and measured the speed of light in water. He found it to be substantially less than the velocity in air. This single experiment settled the century-old uncertainty: Huygens' ideas had been right and Newton's wrong.

Questions

(16-1) **1.** In Galileo's experiment on the velocity of light, suppose that (due to the inevitable human reaction time) there was a half-second delay between the assistant's perception of the flash, and his opening of the shutter. If the two experimenters were, say, 2 miles apart, why did not Galileo conclude the speed of light to be 8 mi/sec? What further experiments would have eliminated (and undoubtedly *did* eliminate) this erroneous conclusion?

2. How far away is the Andromeda Galaxy?

3. The radius of the earth's orbit is 150×10^6 km. About how far ahead of their average schedule will eclipses of a moon of Saturn take place when Earth and Saturn are lined up with the sun, and on the same side of the sun?

4. Roemer's somewhat inaccurate observations led him to believe that the eclipses of Jupiter's satellites were about 11 min ahead of their average schedule when the earth was nearest Jupiter. The diameter of the earth's orbit was at that time thought to be about 180×10^6 mi. What did these values give for the speed of light?

5. In an experiment like that shown in Fig. 16-2 the distance l was 8000 m, the wheels had 360 teeth, and rotated at a speed of 1530 rev/min. What value would this give for the speed of light?

6. If one used a distance $l = 10^4$ m and a 300-tooth wheel in an experiment similar to that shown in Fig. 16-2, how fast would the toothed wheel have to rotate, in rev/min?

(16-2) **7.** A desk is 8 ft below a 950-cp lamp mounted on the ceiling. How many foot-candles of illumination fall on the desk from this lamp?

8. A 750-cp lamp in the ceiling is 6 ft above the surface of a desk. How many foot-candles of illumination does this lamp provide the desk?

9. Noon sunlight may produce as much as 10,000 foot-candles of illumination. The sun is 93×10^6 miles away. Approximately what is the candlepower of the sun?

10. A brilliant arc light suspended 30 ft above a street provides an illumination of 3 foot-candles at a point directly below it. What is the candlepower of the light?

11. A 32-cp standard lamp and a lamp of unknown brightness are 100 cm apart. A grease spot on a piece of paper between them disappears when it is 24 cm from the standard lamp. What is the candlepower of the unknown lamp?

12. A grease-spotted paper is 20 in. from a standard lamp of 60 cp. The spot vanishes when an unknown lamp is 16 in. from the paper, on the opposite side. What is the candlepower of the unknown lamp?

13. A 10-watt lamp and a 40-watt lamp are 1 m apart. A grease-spot photometer sheet placed 29 cm from the 10-watt lamp is equally illiminated on both sides. How does the candlepower per watt of the 40-watt lamp compare with that of the 10-watt lamp?

14. A lamp S and an ordinary 40-watt lamp are 1 m apart, and a grease spot disappears when the paper is 60 cm from S. The 40-watt lamp is replaced by a

100-watt lamp, and the paper must now be placed 44 cm from S to have equal illumination on both sides. How does the candlepower per watt of the 100-watt lamp compare with that of the 40-watt lamp?

(16-3)

15. What color would a red flower appear to be when viewed through a sheet of blue glass?

16. What color would a sheet of white paper appear to be when illuminated by red light?

17. A blue flower is illuminated by sunlight which has passed through a sheet of red glass. What color does the flower appear to be?

18. Seurat, the French pointillist painter, produced color tints, not by mixing pigments, but by placing thousands of dots of different colors side-by-side on his canvas. Discuss how these colors would differ in appearance from those he would have obtained by mixing the same pigments.

(16-4)

19. To an astronaut orbiting above the earth's atmosphere, does the green color of grassy plains appear to be bluer or redder than it does to us on the earth's surface?

20. To an astronaut above the atmosphere, a certain blue star and a certain red star appear equally bright. Do they appear equally bright from the earth's surface? If not, which is the brighter?

(16-5)

21. Small rainbows can often be seen formed by the water droplets from a lawn sprinkler. If such a rainbow resulted from a spray of oil droplets ($n = 1.50$), would the radii of the colored arcs be larger or smaller than those of water droplets, for the primary (i.e., single reflection) bow?

22. In an oildrop rainbow as in Question 21, how would the radii of the colored arcs compare with those formed by water droplets, for the secondary (i.e., double reflection) bow?

(16-6)

23. What is the velocity of light from the hydrogen C line in the glass of Question 43, Chap. 15?

24. What is the velocity of light from the hydrogen F line in the glass of Question 44, Chap. 15?

25. Which travels the faster in glass or water—blue light or yellow light?

26. Which travels the faster in glass or water—yellow light or red light?

chapter / seventeen

The Wave Nature of Light

17-1 Surface Waves

In this chapter (as its title indicates), we shall be investigating the wave nature of light. From a practical point of view, this is not an easy thing to do. The wave nature of light eluded hundreds of intelligent and capable experimenters for many centuries. Even the great Isaac Newton, after considering whether light might actually be a wave phenomenon, finally decided to reject the idea in favor of the old theory that light was composed of particles. So it will be a good idea to first look at some characteristics of wave behavior in a form that is more obvious and easier to observe.

"If you are dropping pebbles into a pond and do not watch the spreading rings, your occupation should be considered as useless," said the fictional Russian philosopher Kuzma Prutkoff. And we can learn much by observing these graceful circles spreading out from the punctured surface of calm water. When one of the waves encounters an obstacle such as, for example, the wall of the pond, it is reflected backward, as shown in Fig. 17-1. The reflected wave looks as if it had been caused

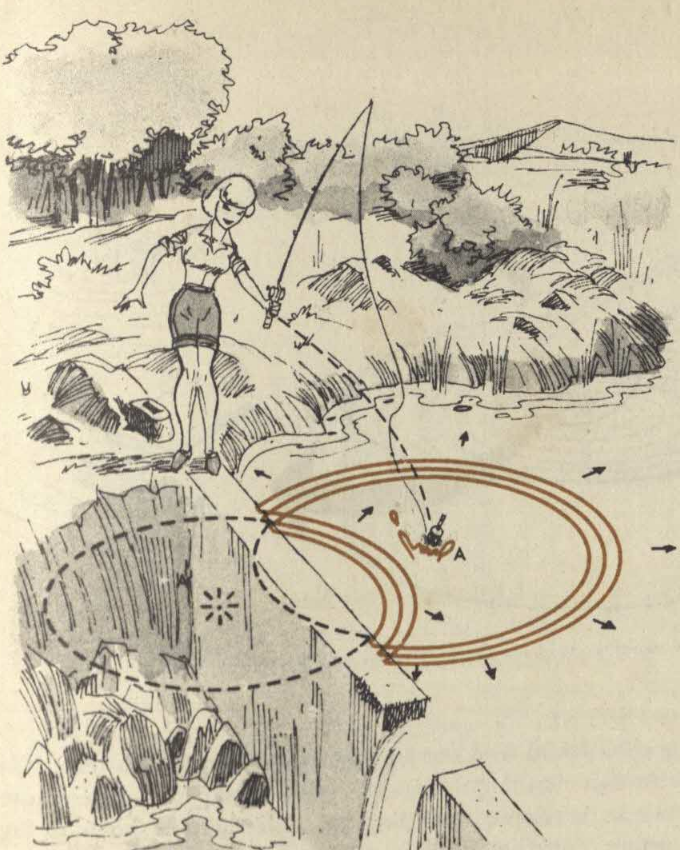


FIG. 17–1 The waves caused by an object thrown into a pond will be reflected from the side of the pond. The reflected waves look as though they originated from an object tossed into an imaginary pond at an exactly opposite and equidistant point on the other side.

by a pebble dropped in the water at an exactly opposite point on the other side of the pond's wall. Thus the wall of the pond acts as a mirror with respect to the surface waves, and, indeed, an optical mirror is based on the same principle except that it reflects light waves instead of water waves.

For a more detailed study of the propagation of surface waves, it is convenient to move from the pond into the laboratory and to generate waves by a more controllable means than tossing pebbles. Figure 17-2 shows a dish filled with water or mercury and an electrically driven vibrating strip which operates on much the same principle as an ordinary electric bell or buzzer. To the vibrating end of the strip we can attach a single needle; or two needles; or a long straight strip—all adjusted to barely touch the surface of the liquid and serve as a source of regular wave trains as they oscillate up and down.

Figure 17-2 shows a series of concentric waves spreading out from a single vibrating point. As the point moves up and down, alternate crests and troughs move out from it with equal speeds in all directions, to form this simple pattern.

If we now replace the single needle with two needles, we should expect

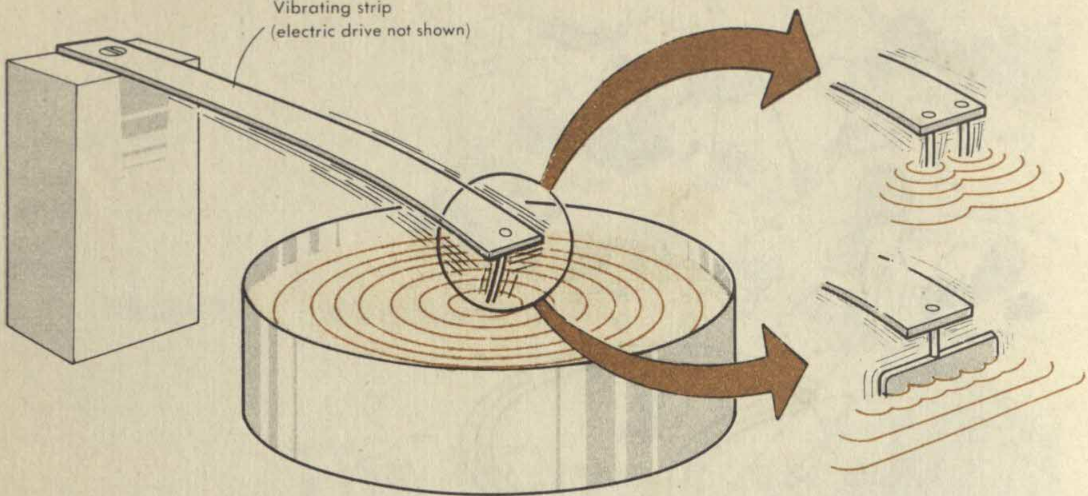


FIG. 17–2 A vibrating strip to generate waves on the surface of a liquid.

that each needle would send out an identical concentric pattern. This is true, but the fact that the two sets of circles overlap each other causes what is known as *interference* between them, with results as shown in Fig. 17-3. The surface of the liquid breaks up into a number of strips marked *C* and *D* on the photograph. Along the strips marked *C* (for *constructive* interference), the crests and troughs are deeper than they would be from either source alone; along the strips marked *D* (for *destructive* interference), the surface is relatively smooth and undisturbed.

An explanation of this behavior is shown in Fig. 17-4, which represents the two sources O_1 and O_2 of Fig. 17-3B, and the line *AB*, drawn anywhere we wish. The point C_0 is equidistant from the two sources,

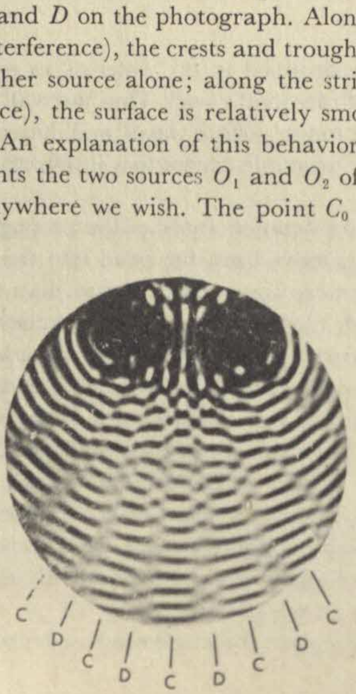


FIG. 17–3 Interference of two sets of concentric waves propagating from two points.

Courtesy The Ealing Corporation.

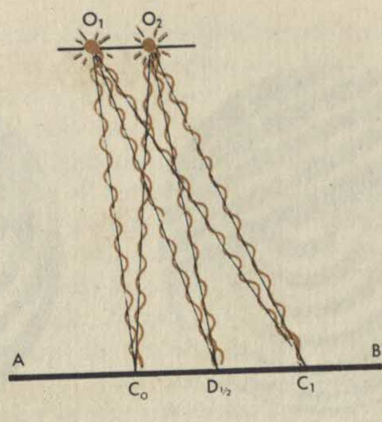


FIG. 17–4 The interference of two identical trains of waves coming from the points O_1 and O_2.

and the waves arrive at C_0 "*in phase.*" That is to say, the crests arrive together and the troughs arrive together, so that their effects add up at this point, and the amplitude of motion of the surface is twice what it would be from either source alone.

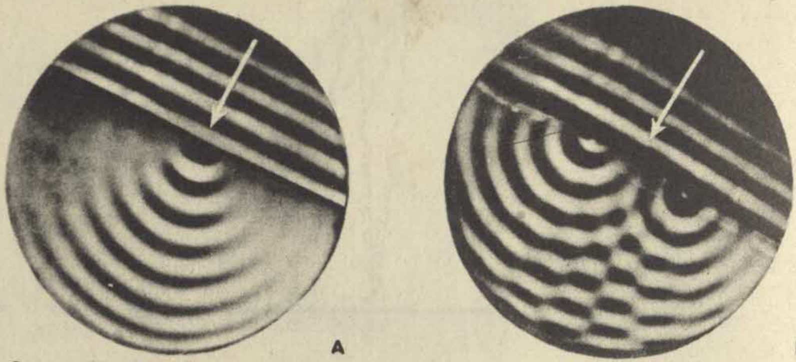
If we go along *AB* a short distance to either side of C_0, we soon come to a point marked $D_{1/2}$, which is just half a wavelength farther from O_1 than it is from O_2. Here a crest from O_1 will arrive just in time to meet a trough from O_2, and vice versa, so that the two impulses will always add up to zero. This is destructive interference, and the surface at $D_{1/2}$ will remain undisturbed.

Farther along *AB* in either direction we shall come to points C_1 and C_2, at which the distances from the two sources will differ by λ, or by 2λ, etc. Wherever the difference is an integral number of wavelengths, the waves will arrive in phase to produce a point of maximum disturbance of the surface. At $D_{3/2}$, $D_{5/2}$, etc., they will arrive exactly out of phase, so that each wave annuls the effect of the other.

If, in Fig. 17-3, we were to draw a number of lines like *AB*, we could plot along each of them similar points of maximum disturbance (antinodes) and points of zero or minimum disturbance (nodes). Then, by connecting corresponding points, the nodal lines and antinodal lines could be traced out to exactly follow their locations as shown in the photograph. (Students with some analytical geometry will recognize that these lines are hyperbolas.)

17-2 Huygens' Principle

We could replace the vibrating needles on the wave generator by a long straight strip (Fig. 17-2), which would generate a series of straight-line waves propagating across the tank. (We could also generate waves which were practically straight lines, by placing a single vibrating needle at a great distance away. But since the waves get weaker with increasing distance, this would introduce too many practical difficulties.)



Courtesy The Ealing Corporation.

FIG. 17-5 The diffraction of a plane wave passing through a small opening in a breakwater (A); and the interference between two such wave patterns formed by two openings (B).

Now, if in the way of these straight waves we place a barrier with a single relatively narrow opening in it, the waves will pass through the opening and produce a pattern, beyond the barrier, similar to the pattern produced by a single oscillating needle (Fig. 17-5A). If the barrier, or "breakwater," has two openings, the pattern beyond it is similar to that produced by a pair of oscillating points (Fig. 17-5B).

In the paragraph above, we mentioned the wave "passing through" the opening in the breakwater. In a way, this statement is true, of course, but we shall do better to consider the matter from the point of view of Huygens. Looking up from the lower side of Fig. 17-5A, we could see the water in the slot of the breakwater oscillate up and down as the plane waves on the other side impinge against it. The disturbance caused by this oscillating water spreads out from the slot in concentric ripples, just as it does from the disturbance under an oscillating needle. In a similar way, the water rising and falling in the two slots of Fig. 17-5B produces two independent point sources of interfering waves.

These examples are illustrations of *Huygens' principle*, which says that ***every point along a wave front serves as a source from which new waves spread out.*** This principle is true, even though we may not always have a convenient breakwater to shut off all its effects except those from one or two isolated points along a wave front. The spreading of concentric ripples on an unobstructed surface, or the advance of plane waves, results from the mutual interference, both backward and forward, among the wavelets sent out simultaneously from an infinite number of points along every wave. Huygens' principle is of great importance in theoretical optics and has an important bearing on our study of light waves.

17-3 Interference of Light Waves

We have already described interference phenomena in connection with the waves on the surface of a liquid, and the same reasoning can be applied to the interaction of two light waves.

But how can we synchronize the vibrations of two light sources? Visible light, as we shall soon see, is composed of an enormous number of trains of transverse waves, called light *quanta*, or *photons*. Light is emitted a photon at a time by separate individual atoms. The photons have no synchronization with one another, and if we tried to observe interference effects between light from two different sources we could see none at all. *The only way to observe interference is to let each photon interfere with itself*; in this way, perfect synchronization is guaranteed for all photons. It is very simple to do this, in a manner analogous to the two-opening breakwater shown previously for surface waves in Fig. 17-5B. Figure 17-6 shows a source of monochromatic light (i.e., light that is all of the same wavelength) illuminating a slit, from which light spreads out to strike two other slits, S_1 and S_2. Each photon that strikes the first slit may be considered to strike S_1 and S_2 simultaneously, and from S_1 and S_2 waves fan out that are accurately in phase with each other. For the sake of clarity, waves have been graphically indicated as traveling along straight rays to points A, B, and C on a screen or a piece of photographic film. Point A is equidistant from S_1 and S_2, so that the waves starting together from the two slits arrive exactly in phase at A, and in this manner reinforce each other to produce a bright illumination. Point B on the drawing was selected so that it is half of a wavelength,

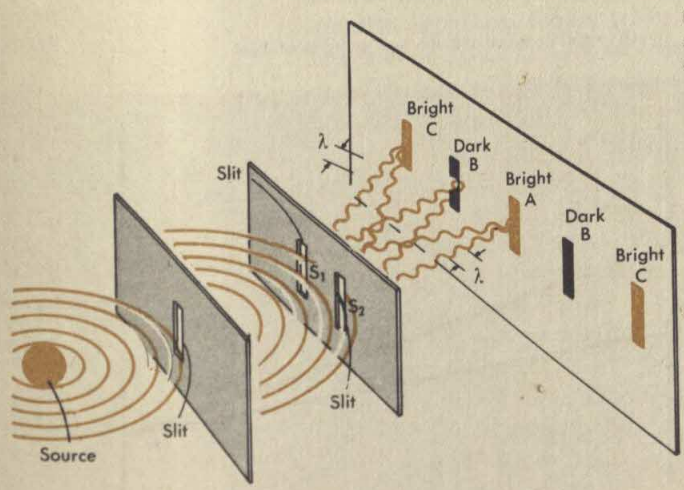


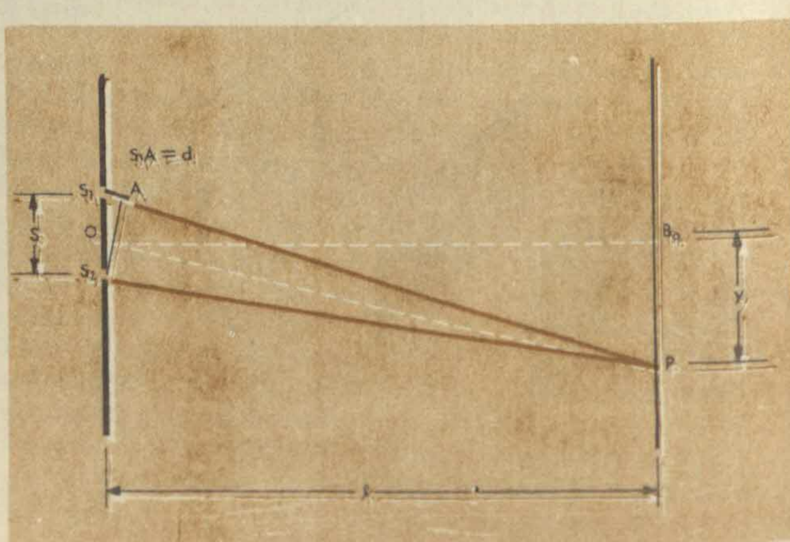
FIG. 17–6 Light and dark bands on a screen, formed by the interference of light coming from two closely spaced slits.

$\lambda/2$, farther from S_2 than from S_1. Waves starting out in phase will thus arrive at B exactly out of phase and will annul each other, so that B remains dark. Point C is just λ farther from S_2 than it is from S_1, so when the waves arrive at C they are shifted one wavelength with respect to each other. This puts them in phase, and C is bright.

Such reasoning can be continued, and we can make a general statement that any point which is an odd number of half-wavelengths closer to one slit than the other will be dark; any point which is the same distance from both slits, or from which the distances to the slits differ by an integral number of whole wavelengths, will be of maximum brightness. The English scientist, scholar, and engineer Thomas Young (1773–1829, the same man whose name was given to Young's modulus), performed experiments of this sort in 1800, and his work, together with that of his contemporary Augustin Fresnel (1788–1827) in France, who built on the previous work of Christian Huygens, brought the corpuscular theory of light to an end. (This is not quite true; in recent years, light has been found to have some aspects resembling those of particles, but in an entirely different sense than that which Newton had in mind. We shall look into this a few chapters further along.)

Interference methods provide a direct way of measuring the wavelengths of light of different colors. If the slits are illuminated with blue light, the pattern of light and dark bands will be much more closely spaced that if they are illuminated with red light. This demonstrates that blue light has a shorter wavelength than red light. The shortest wavelength visible to human eyes is about 4×10^{-5} cm and is beyond the blue color into the violet. The longest wavelength that human beings

FIG. 17-7 terference patterns and the wavelength of light.



can see, a deep red, is about 7.5×10^{-5} cm. Wavelengths are often given in *Ångstrom units* (Å). One Ångstrom unit is 10^{-8} centimeter, so we can say that the visible wavelength range is from 4000 to 7500 Å.

As an example of how a pair of slits might be used to determine the wavelength of some particular color of light, let us turn to Fig. 17-7 for the geometry of the arrangement. The two slits S_1 and S_2 are separated by a distance s. They are illuminated from a single slit or other small source of monochromatic light somewhere off to the left of the drawing and which is not shown. If we assume that the incoming waves strike S_1 and S_2 simultaneously, we may mark point B_0 on the screen, directly opposite point O, midway between the slits. Therefore B_0 is equidistant from S_1 and S_2, and no matter what wavelength strikes the slits, it will generate wave trains from S_1 and S_2 which will arrive at B_0 in phase and therefore reinforce each other to produce a bright line there.

Now let us move down (or up) the screen to point P, a distance y from B_0. As we do so, we mark point A on the line S_1P, so that $AP = S_2P$. The path from S_1 to P is longer than the path from S_2 to P by the distance S_1A, which we can refer to by the single letter d. If d is equal to any *integral* number of wavelengths, waves from S_1 and S_2 will arrive *in phase*, and P will mark a location of maximum brightness. If d is equal to any odd number of half-wavelengths ($\frac{1}{2}$, $\frac{3}{2}$, $\frac{5}{2}$, etc.), the waves will arrive exactly *out of phase*, and P will mark a dark line on the screen.

As Fig. 17-7 has been drawn, OB_0 is perpendicular to S_1S_2; OP is perpendicular to AS_2. Elementary geometry (or even elementary common sense reasoning, on which geometry is based) then shows that angles S_1S_2A and POB_0 are therefore equal. The careful geometer will see that triangle S_2S_1A and triangle OB_0P are not quite similar, but he will also agree that the deviation from exact similarity is too small to bother about when s is long in comparison with d, or l long in comparison with y, as in nearly always the case. Accepting this similarity of triangles, then, we have

$$\frac{d}{s} = \frac{y}{l}.$$

As an example, let us suppose that we have a pair of slits whose separation is 6.1×10^{-3} cm and that they are mounted 25 cm in front of a photographic film. The slits are illuminated by monochromatic light, and after a proper exposure, the film shows a series of bright bands spaced 0.21 cm part. Although the bands are evenly spaced and it makes no difference which pair we choose, let us follow Fig. 17-7, and choose the central bright band B_0 and one of its nearest bright-line neighbors B_1, which will be just $y = 0.21$ cm away. Since B_1 is bright, d is exactly one wavelength, and

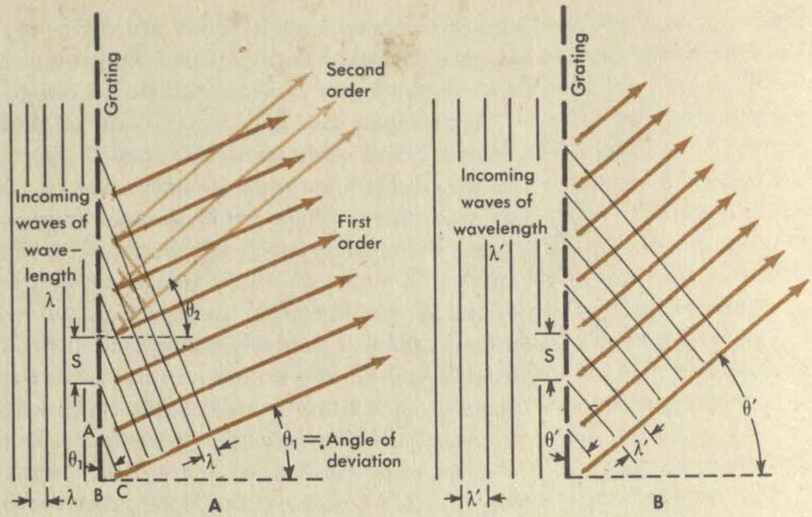


FIG. 17-8 The direction in which a grating reinforces light that falls on it depends on the wavelength of the light. The longer the wavelength, the greater the angle of deviation—the reverse of the behavior of a prism.

$$\lambda = d = \frac{sy}{l} = \frac{6.1 \times 10^{-3} \times 0.21}{25} = 5.1 \times 10^{-5}\,\text{cm}$$

or

$$\lambda = 5.1 \times 10^{-5} \times 10^{8} = 5100\,\text{Å}.$$

17-4 Optical Gratings

The idea of the interference of water waves can be extended to a very long breakwater with many openings in it. This is analogous to the effect of an optical "grating" on light waves. We can make a *transmission grating* from a piece of glass on which we have scratched a large number of fine parallel grooves. When placed in a beam of light, the glass will transmit, i.e., let pass through, only the light that falls on the smooth unscratched strips between the grooves. A *reflection grating* works in a similar manner by reflecting light from the smooth sides of accurately shaped grooves in the surface of a metal mirror. With modern techniques, we can make gratings of several thousand scratches per millimeter. Such original gratings are very expensive and are ordinarily used for only the most important and delicate research. It is possible to take impressions in plastic from an original grating. These may be nearly as accurate as the original, much cheaper, and satisfactory for many purposes.

Figure 17-8A shows the wave fronts of a beam of light of wavelength λ falling on a transmission grating. The *grating spacing*, which is the distance between lines or scratches, is *s*. By Huygens' principle, each line

of the grating will serve as a source of new waves spreading out to the right of the grating. For the particular wavelength λ, there will be a particular direction, deviated by the angle θ_1, in which the waves from each slit will be exactly one wavelength ahead or behind the waves originating in the adjacent slits. Accordingly, *in this direction*, all the waves from all the slits will be in phase with each other.

The direction θ_1, in which the light from each line is just one wavelength λ ahead or behind the waves from adjacent slits, is the direction of the *first-order* spectral line of wavelength λ. In Fig. 17-8A, we can see another direction θ_2, in which the waves from each grating line are 2λ ahead or behind the adjacent waves. This is the direction of the *second-order* spectral line of wavelength λ.

The small triangle marked ABC on Fig. 17-8A enables us to derive a simple relationship that determines the angle of deviation θ for any wavelength. Angle BAC equals θ_1, and since the direction of propagation BC is perpendicular to the wave front AC, angle ACB is a right angle. We can thus write

$$\sin \theta_1 = \frac{BC}{AB}.$$

But $BC = \lambda$, the wavelength of the light, and AB is s, the grating spacing. Accordingly,

$$\sin \theta_1 = \frac{\lambda}{s}.$$

For the second-order direction, it is apparent that

$$\sin \theta_2 = \frac{2\lambda}{s}.$$

Figure 17-8B shows light of longer wavelength striking the same grating. Analysis will show that since λ' is greater than λ, θ' must be greater than θ. In other words, in a grating spectroscope, the longer wavelengths are deviated more than the shorter wavelengths. (The opposite happens when light passes through a prism.)

Most spectroscopes now use gratings rather than prisms for the investigation of the components of light present in its spectrum. Gratings serve the same purpose as prisms in deflecting light of different colors, or wavelengths, through different angles which can then be accurately measured. The arrangement of Fig. 17-9 is the same as that for the prism spectroscope shown in Fig. 15-17, except that the prism is replaced by a grating whose plane is perpendicular to the rays from the collimator.

The wave nature of light is also revealed when light passes through a very small opening that has dimensions comparable to the wavelength of light. If the opening is very much larger than the wavelength, light passing through it will make a spot of light the same shape as the

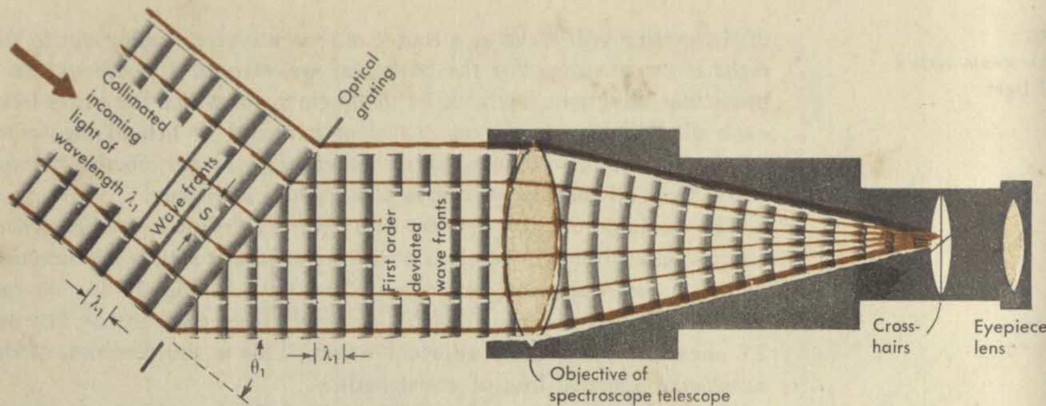


FIG. 17-9 Schematic diagram of a grating spectroscope. The first-order deviation of the incoming monochromatic radiation is focused on the crosshairs of the spectroscope telescope.

opening on a screen placed behind it. However, with an opening whose dimensions are comparable to the wavelength of light, the light will be scattered (diffracted), and all that will be seen on the screen will be a diffused, luminous spot composed of concentric light and dark rings, from which we cannot conclude anything very definitely about the shape or size of the opening. This phenomenon, known as *diffraction of light*, places a lower limit on the size of small objects that can be seen or photographed by using visible light. In fact, light waves cannot produce a picture of objects that are only the length of a lightwave for the same reason that a painter cannot paint a miniature portrait using a 2-inch brush.

17-5 Polarization of Light

We saw in an earlier chapter that there are two possible means of wave propagation: (1) longitudinal waves in which the motion of individual particles is along the line of propagation and (2) transverse waves in which the motion is perpendicular to the line of propagation.

Which kind of wave motion do we have in the case of light? The important difference between longitudinal and transverse waves is that the latter can be *polarized*. To understand this important notion, let us look at a wave along the direction of its propagation, as shown in Fig. 17-10. In the case of longitudinal waves (Fig. 17-10A), the motion of the particles takes place perpendicular to the surface of the paper and will not be noticeable from the direction we are observing it. In the case of transverse waves—Fig. 17-10B and C—the motion of the particles is in the plane of the paper and easily observable in that projection. We call the transverse wave *unpolarized* if the motion of the particles

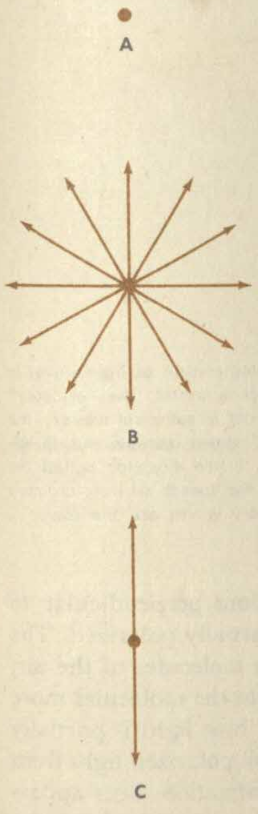
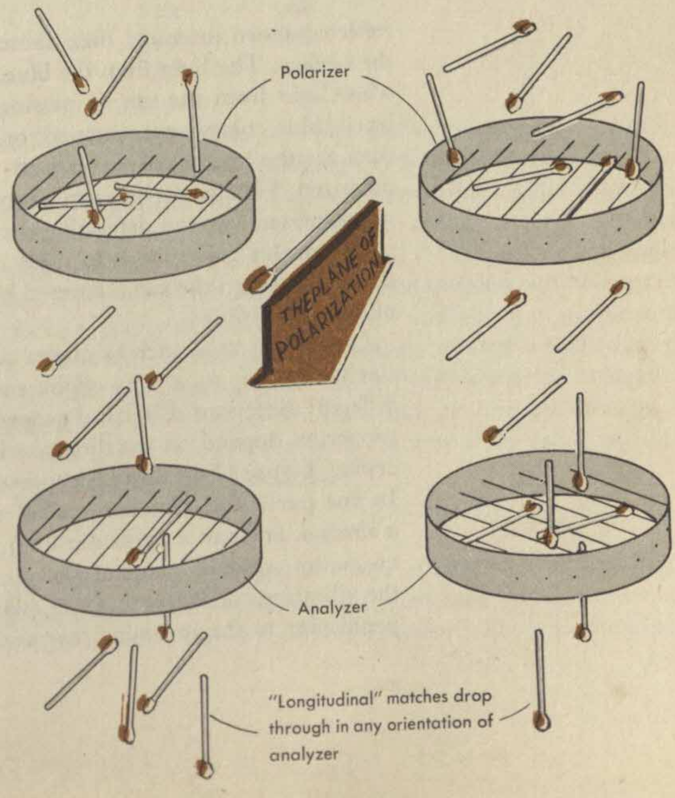


FIG. 17–10 Cross sections of different sorts of waves, looking along the direction of their propagation.

takes place in all possible directions (Fig. 17-10B); if the motion is only in one direction (Fig. 17-10C), the wave is *polarized*. The notion of polarization can be clarified by the analogy given in Fig. 17-11. Suppose we have a sieve, made of a set of parallel wires without crosswires, and we drop matches on it in such a way that the falling matches remain horizontal but may have different orientations in the horizontal plane. It is clear that only those matches parallel to the sieve wires will pass through. If under this "match polarizer" we place a similar "match analyzer," the "beam of matches" will pass through only if the wires in the lower sieve run parallel to those of the upper sieve. We can extend this analogy to the case of longitudinal waves by dropping the matches in a vertical position. In this case, all the matches will pass through, regardless of the relative positions of the two sieves.

To our eyes, the plane of vibration of light waves makes no difference; light that has been polarized so that the vibration is all in a single plane (Fig. 17-10C) looks exactly the same as ordinary light that consists of myriads of photons whose planes of vibration are aligned in every possible direction (Fig. 17-10B). Much of the light we see is at least partially polarized, because whenever light is reflected from a smooth *nonmetallic* surface, the waves with vibrations parallel to the surface are

FIG. 17–11 "Polarized" matches.



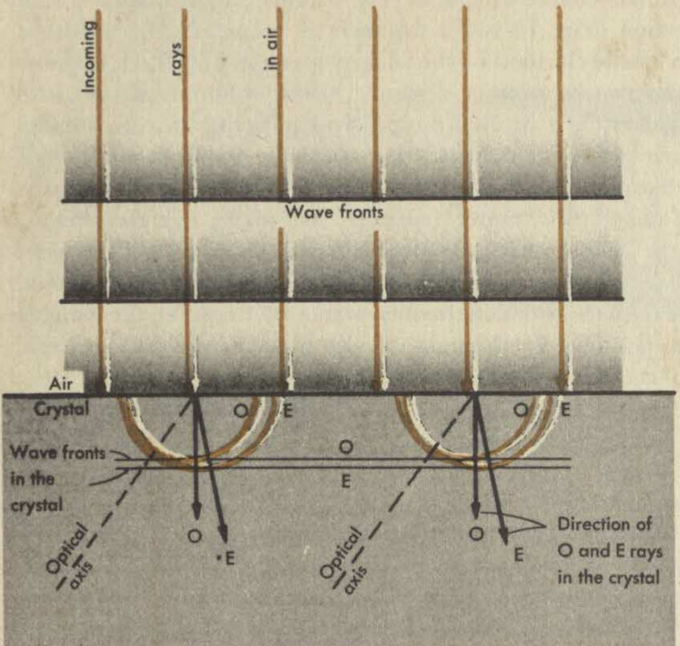


FIG. 17–12 Transmission of light waves in a doubly refracting crystal. The "ordinary" waves spread out in spherical waves; the "extraordinary" waves spread out in elliptical waves. In the direction called the "optical axis" the speeds of both ordinary and extraordinary waves are the same.

reflected more intensely than those with vibrations perpendicular to the surface. The light from the blue sky is also partially polarized. The white light from the sun, in passing through the molecules of the air, has its blue component scattered, or reflected, from the molecules more than are the longer red waves, and this scattered blue light is partially polarized. The eyes of bees can not only distinguish polarized light from nonpolarized but can determine the plane of polarization. Bees apparently use this ability to help them navigate back to the hive and can do so even when the sun is covered by clouds, as long as there is a patch of blue sky visible.

Certain crystals, such as quartz and calcite, as well as many others (including ice, to a very slight extent) have different properties in different directions. Electrical properties, heat conductivity, and optical properties depend on the direction in which they are measured in the crystal. Figure 17-12 shows the spreading of light waves in such a crystal. In one particular direction, called the *optical axis* ("axis" here means a *direction*, and not a particular line), all light travels at the same speed, no matter what its plane of vibration. In Fig. 17-12, we may imagine the vibrations of the light waves (whatever they may be) as being perpendicular to the incoming rays and oriented in all possible directions

around the rays. These vibrations can be broken down into components parallel to the optical axis and components perpendicular to the optical axis. Experimentation has showed that those vibration components perpendicular to the optical axis behave in a perfectly ordinary way; in using Huygens' principle, we can show them spreading out with the same velocity in all directions, so that the wave fronts we imagine will be little spheres. The components parallel to the optical axis, however, behave in a very extraordinary way by traveling at different velocities in different directions. These Huygens' wavelets spread out, not in spheres, but in ellipsoids, as shown. We thus have an incoming ray broken into two parts: the ordinary ray, which obeys Snell's law; and the extraordinary ray, which does not. (Remember that in deriving Snell's law, we assumed the velocity of propagation to be the same in all directions; we could not expect it to apply in its simple form to materials for which this assumption is not true.)

We have defined the index of refraction n to be the ratio of the speeds of light in air (or vacuum) and in the material. A doubly refracting crystal will have two indices of refraction: n_E for the extraordinary ray and n_O for the ordinary ray. Indices of refraction for the ordinary ray and the extraordinary ray are given here for a few commonly used materials:

	n_O	n_E
Calcite	1.658	1.486
Quartz	1.544	1.553
Tourmaline	1.637	1.619

Tourmaline has the property of being relatively opaque to the ordinary ray, so that only the extraordinary ray is transmitted through it. Thus, if a ray of unpolarized light falls on a tourmaline crystal in a direction perpendicular to its optical axis, the light coming through the crystal will emerge completely polarized, i.e., with all the vibrations in a single plane that is parallel to the optical axis. Iodo-quinine sulfate crystals have this same property of absorbing one ray and being transparent to the other. "Polaroid" is composed of billions of tiny needle-shaped iodo-quinine sulfate crystals lined up with their optical axes parallel and imbedded in a sheet of plastic. Light passing through a sheet of Polaroid is almost completely polarized, the vibrations being parallel to the length of the imbedded crystals.

17-6 The Nature of Light Waves

Experiments showing the interference effects of light showed that light must have the characteristics of waves; the phenomenon of polarization showed further that the waves must be transverse. Early in the nine-

teenth century, the nature of these waves was still a complete puzzle to scientists.

Considering light as transverse waves propagating through space, they logically assumed that there must be some medium through which these waves propagated. Since light travels easily through empty space, this hypothetical medium had to be assumed to fill all space and to also penetrate the interior of material bodies. It was called "light ether," "world ether," or more simply just "ether." It had to be rigid and substantial enough to transmit transverse light waves through enormous distances from faraway stars; yet it also had to be yielding enough to let planets and satellites pass through it at high speed with no measureable friction. No satisfactory way was ever found to reconcile these conflicting requirements.

In 1864, the great mathematical physicist James Clerk Maxwell published a paper dealing with the behavior of electric and magnetic fields. In this paper, he showed that under the proper circumstances a combination of electric and magnetic fields (or more briefly, an electromagnetic field) would propagate through space—and at a speed of 3×10^{10} cm/sec. It is important to remember that Maxwell's paper was not dealing with light; the figure 3×10^{10} cm/sec was derived from the ratio of certain electrical and magnetic units. The coincidence of this figure with the velocity of light was too close to be accepted as a mere coincidence, and most scientists accepted it as a good indication that light was an electromagnetic wave. An indication is not a proof, however, and in 1888 a German physicist, Heinrich Hertz, set out to translate Maxwell's theoretical equations into a practical demonstration. Let us review the basic principles behind Hertz's work.

If we take a capacitor and give its plates opposite electric charges, we shall have had to do work in some way to pull electrons off the positively charged plate and to force excess electrons on the already negatively charged plate. If we look to see where this work has gone, we shall find it stored in the electric field that now exists between the plates. Suppose we now connect these capacitor plates to a solenoid (Fig. 17-13A). A current will begin to flow from one plate through the solenoid to the other plate, and as the charge on the plates thus becomes less, the electric field between them will also decrease. What happens to the energy stored in the electric field as the field becomes weaker? Of course the answer is that we can find this lost energy now stored in the magnetic field caused by the current through the solenoid (Fig. 17-13B). In Fig. 17-13C, we see the situation when the capacitor is completely discharged: the electric field has vanished, and all the energy is in the magnetic field. There is no difference in charge to keep the current flowing, and if this were the whole story, the current would stop. But Le Chatelier's principle now goes to work to maintain the status quo, i.e., to prevent the current from stopping. The magnetic

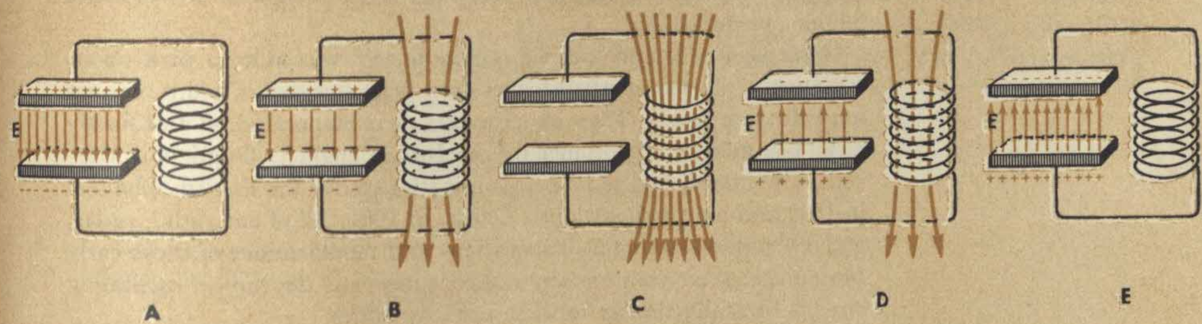


FIG. 17-13 Electromagnetic oscillations between a capacitor and a solenoid.

field delivers its stored energy back into the wire to keep the current flowing (Fig. 17-13D), and the charge begins to build up on the capacitor in the opposite direction. Finally, when the magnetic field has been reduced to zero, the current stops, and we have the same situation we had in the beginning (Fig. 17-13E), except that the sign of the charges has been reversed. Now, of course, the whole cycle will repeat itself in the reverse direction, and again and again, and so on.

This oscillating electric current is analogous to the behavior of a pendulum. We start it by pulling the pendulum aside and giving it potential energy (we give the conductors the energy stored in the electric field). As it swings down, this potential energy is converted into kinetic energy (the energy of the electric field is now in the magnetic field). The pendulum swings over to the other side, and its kinetic energy is again converted to potential energy (the magnetic field has vanished, and its energy appears in the electric field again). And, just as the pendulum's period can be changed by changing its length, the period of electric oscillation in our system can be changed by changing either the capacitor or the solenoid.

The pendulum, of course, will not swing forever; it gradually loses its energy by friction. The electrons moving back and forth in our oscillating electrical system encounter resistance, too, in the wire through which they flow, and unless its energy were periodically replenished, it would also come to a stop.

This analogy can be carried a step further. If we suspend our pendulum from a clothesline instead of from a rigid support, some of the pendulum's energy will go to moving the clothesline back and forth

and will be dissipated in waves traveling out along the line in both directions. Part of the energy of our oscillating electric circuit will in a somewhat similar way go into creating "electromagnetic waves" that radiate out into space.

Hertz used a simple electric oscillator and was able to pick up its radiation of electromagnetic waves on a primitive sort of radio receiver some distance away. These experiments were elaborated on and finally given commercial importance by the Italian engineer Guglielmo Marconi, who established radio communication across the English Channel in 1899 and across the Atlantic Ocean in 1901. All of our radio, radar, and TV today are merely elaborations and modifications of those early beginnings; their transmission and reception still depend on oscillating circuits basically similar to those used by Hertz.

17-7 Electromagnetic Spectrum

Radio waves and light waves may at first glance seem to be very different, but actually they are the same thing—electromagnetic waves which differ only in their frequencies. We now have explored an enormous spectrum of electromagnetic radiation, over a wide range of frequencies and wavelengths—but all having the same velocity in space: 3×10^{10} cm/sec, as predicted by Maxwell in 1864.

Radiation visible to our eyes occupies only a small part of this great range. The reddest light we can see has a wavelength of about 7.5×10^{-5} cm ($f = 4 \times 10^{14}$ hz*). The other, blue-violet end of the visible range occurs at a wavelength of about 4.0×10^{-5} cm ($f = 7.5 \times 10^{14}$ hz).

Table 17-1 ELECTROMAGNETIC RADIATION

Name	Frequency Range	How Produced
60-cycle	60 hz	the weak radiation from our alternating-current circuits
Radio, radar, and TV	10^{4}–10^{10} hz	oscillating electric circuits
Microwaves	10^{9}–10^{12} hz	oscillating currents in special vacuum tubes
Infrared	10^{11}–4×10^{14} hz	outer electrons in atoms and molecules
Visible	4×10^{14}–8×10^{14} hz	outer electrons in atoms
Ultraviolet	8×10^{14}–10^{17} hz	outer electrons in atoms
X rays	10^{15}–10^{20} hz	inner electrons in atoms, and sudden deceleration of high-energy free electrons
Gamma rays	10^{19}–10^{24} hz	nuclei of atoms, and sudden deceleration of high-energy particles from accelators

* The *hertz*—abbreviated hz—means "cycles or vibrations per second." Its use not only saves time and trouble, but honors a great scientist.

Nineteenth-century experimenters soon realized that the spectra produced by their prisms and gratings did not come to an end merely because we could not see them. Beyond the violet (*ultraviolet*) were radiations that could, among other effects, register on a photographic plate. Beyond the red end of the visible range (*infrared*) were radiations easily detectable by their ability to heat sensitive thermometers. The work of Hertz extended this range still more, as did the discovery of X rays and gamma radiation. Table 17-1 gives some indication of the range of electromagnetic radiation.

Questions

(17-1)

1. Two needles 10 cm apart vibrate with a frequency of 15 cycles/sec and generate waves on a liquid surface. The velocity of the waves is 25 cm/sec. (a) What is the wavelength of the generated waves? (b) How many antinodal points (of maximum disturbance) will there be on the line joining the two vibrating points? (c) How many nodal points (of zero or minimum disturbance?)

2. A liquid surface is disturbed by two vibrating points 8 cm apart, which oscillate at a frequency of 12 cycles/sec. The liquid surface transmits these waves at a speed of 30 cm/sec. (a) What is the wavelength of the waves? (b) How many points of maximum disturbance (antinodes) are there on the line joining the vibrating points? (c) How many points of zero, or minimum disturbance (nodes)?

3. Are any of the nodal points located in Question 1 actually points of zero disturbance? (Bear in mind that the amplitude of the waves diminishes with distance.)

4. On a line like AB in Fig. 17-4, is the disturbance actually zero at any of the nodal points? (Bear in mind that the amplitude of the waves diminishes with distance.) Is the resultant disturbance less at $D_{3/2}$ or at $D_{5/2}$?

(17-3)

5. A pair of narrow parallel slits are 0.03 cm apart, and 20 cm in front of a photographic film. If the slits are illuminated with monochromatic light of $\lambda = 5710$ Å from another narrow slit, how far apart will the resulting bands be on the film?

6. Thirty centimeters in front of a photographic film are a pair of narrow slits, 0.04 cm apart. These slits are illuminated by light of $\lambda = 5893$ Å from another slit. How far apart are the bright bands on the film?

7. The equipment of Question 5 is illuminated by radiation of unknown λ, and the resulting bands (often called "fringes") on the film are 1.70×10^{-2} cm apart. What is the λ of the radiation? Would it be visible to the eye?

8. The experimental setup of Question 6 is illuminated by monochromatic radiation of unknown wavelength, and the bands on the film are found to be 0.7 mm apart. What is the wavelength of the radiation? Would it be visible to the eye?

(17-4)

9. A grating spectrograph is used to examine light of $\lambda = 6200$ Å. The grating has 5000 lines/cm. The "zero order" image will of course be seen when the collimator and telescope are in line. From this position, (a) through what angle

must the telescope be turned to see the first-order line? (b) Is it possible to see the second-order line? If so, through what angle must the telescope be turned? (c) Same as (b), for third-order.

10. A grating with 4000 lines/cm is used in a spectrograph to examine light of $\lambda = 4960$ Å. (a) Through what angle from its "zero-order" position must the spectroscope be rotated to see the first-order line? (b) the second-order line? (c) the third-order line?

11. A spectrograph using a grating of 12,000 lines/inch must be turned through 17° to put the crosshairs on the first-order line of an unknown radiation. What is the wavelength of the radiation?

12. A grating ruled with 13,400 lines/inch is used to examine radiation of unknown wavelength. The spectroscope is turned through 20° to bring its first-order image on the crosshairs. What is the wavelength of the radiation?

13. The "C" line is a monochromatic component of the light emitted by incandescent hydrogen gas, and has a wavelength of 6563 Å. (a) What is the highest order C line that can be observed with a spectroscope using a grating with 6000 lines/cm? (b) Will this particular line appear to be of the same color in each order?

14. One component of the light emitted by incandescent hydrogen is the F line, $\lambda = 4861$ Å. (a) What is the highest order F line than can be seen through a spectroscope using a grating of 5600 lines/cm? (b) Will this line appear to be of the same color in each order?

(17-5) **15.** Given a small piece of Polaroid, how could you use it to determine whether a light beam was polarized or not?

16. Polaroid sunglasses reduce "reflected glare." Explain how this is done.

17. It has been suggested that the dangers of night driving could be reduced by reducing the glare of oncoming headlights. Can you propose how this might be done with sheets of Polaroid? (Remember that you must be able to see objects properly illuminated by your own headlights.)

18. Two pieces of Polaroid with their axes at 90° almost completely block the passage of any light through them. What will be the effect of slipping another piece of Polaroid between the first two, with its axis oriented at 45°?

19. The faces of a parallel-sided plate of calcite are parallel to the optical axis, and the plate is 1 mm thick. A normally incident ray has a wavelength of 6000 Å in air. (a) What are the wavelengths of the ordinary and of the extraordinary components in the calcite? (b) How many λ_o and λ_E are there in the plate? (c) When 0 and E emerge into the air again, how many wavelengths has the E component gained on the 0 component?

20. The faces of a parallel-sided plate of quartz are parallel to the optical axis of the crystal. (a) What is the thinnest possible plate that would serve to put the ordinary and extraordinary components of a light beam ($\lambda = 5890$ Å) a half-wave apart on their exit? (b) What multiples of this thickness would give the same result?

(17-7) **21.** If the frequency range shown in Table 17-1 were plotted on a linear scale from 0 to 10^{24} hz on a graph 1 km long, how long a piece of this graph would represent visible light?

chapter / eighteen

The Special Theory of Relativity

18-1 The Paradox of the Ether

When the wave nature of light was definitely proved, physicists had a great deal of trouble in their attempts to describe the mechanical properties of the hypothetical "ether" through which the light waves were supposedly propagating. Indeed, since the phenomenon of polarization of light proved beyond any doubt that the scientists were dealing here with transverse vibrations, they were forced to consider the ether as some kind of rigid material presenting resistance to sidewise deformations (shear). But if the entire space of the universe were filled with this hypothetical rigid material carrying light waves emitted by the sun and the stars, how could the celestial bodies, and in particular our earth, move through space without any apparent resistance?

Escape from this apparent contradiction was sought in the assumption that ether had properties similar to those of plastic materials, such as sealing wax, which behave as solids under the influence of strong forces acting over a short period of time, but flow as liquids when acted upon by weak but persistent forces such as their own weight. It was argued

that in the case of propagating light waves, where the force changes its direction 10^{15} times per second, ether could show the properties of an elastic solid; and that on the other hand, it could flow as a perfect fluid around the bodies of the planets of the solar system since in this case the rotation periods are measured in years. But this "explanation" of the unusual properties of the light-carrying medium remained nothing but words, and no consistent theory of its mechanical behavior was ever worked out. Difficulties increased when the development of the electromagnetic theory of light caused scientists to consider electric and magnetic fields as strains and stresses in the world ether. Indeed, if this substance was supposed to behave as a perfect fluid in response to forces acting over long periods of time, how could the static electric and magnetic fields surrounding electrically charged conductors and permanent magnets be ascribed to its elastic deformations?

The contradictions piling up in the problem of the mechanical properties of the world ether heralded the end of the old classical physics and opened the era of modern physics with its very unconventional, and, at first sight, very strange ideas. It was the first time that physicists realized that there are things that cannot be described in terms of the familiar properties of ordinary material bodies such as solids, liquids, or gases. And, indeed, why should whatever it is that is responsible for electric and magnetic fields have the properties of ordinary matter? We know, in fact, that all the familiar properties of ordinary matter, such as elasticity, fluidity, compressibility, and so on, are due to electric and magnetic interactions between the atoms from which the matter is formed. In particular, the elastic properties of solids, such as their resistance to bending or shear, can be derived mathematically from their atomic structure, and the hardness of a diamond as contrasted to the softness of graphite is a direct result of the different arrangement of carbon atoms forming these two materials. Unless one ascribed atomic structure to the ether itself, there was no reason at all to expect that it would have the properties of ordinary solid materials. In fact, assuming the existence of this hypothetical all-penetrating medium, one could just as well ascribe to it any kind of unusual properties that might be needed, provided these properties would lead to a correct interpretation of the known facts concerning the propagation of light and electromagnetic phenomena in general.

18-2 Ether Wind?

The acid test of the ether hypothesis came late in the nineteenth century and resulted in a complete turnabout of our ideas concerning the nature of light waves and electromagnetic fields.

If it were true that light waves propagate through a jelly-like ether which fills universal space, we should be able to measure our motion

through space by observing the effect of that motion on the velocity of light. In fact, since the earth moves in its orbit at a speed of 30 km/sec, we would experience an "ether wind" blowing in the direction opposite to our motion in the very same way that a speeding motorcyclist experiences a strong "air wind" blowing into his face. Light waves propagating in the direction of that "ether wind" would move faster, being helped by the motion of the medium, while those propagating in the opposite direction would be slowed down. In the year 1887, the American physicist A. A. Michelson (1852–1931) carried out an experiment that was expected to demonstrate the effect of the earth's motion on the velocity of light as measured on its surface. Instead of measuring the velocity of light in two opposite directions, Michelson found it more convenient to compare it in two mutually perpendicular directions. In order to understand Michelson's scheme, let us consider a Mississippi steamboat running between St. Louis and Memphis, some 300 miles apart. Sailing downstream, the boat moves faster, since it is helped by the current, whereas on the way back, it is correspondingly retarded. Does the gain in time one way compensate for the loss in time the other? Although it may appear at first sight that it does, this conclusion is not true. Let us do some simple arithmetic, assuming that the boat's speed (in still water) is 30 mi/hr and that the velocity of the stream is 3 mi/hr. Sailing upstream and downstream, the boat will have the velocities, relative to the shore, of 27 and 33 mi/hr. The time necessary for the round trip will apparently be

$$\frac{300}{27} + \frac{300}{33} = 11.11 + 9.09 = 20.20 \text{ hr}$$

which is 1 percent more than it would be in still water. The closer the velocity of the stream is to the velocity of the boat, the longer is the time necessary for the round trip; and if the two velocities are equal, the boat will never return!

Let us consider now the problem of how to move across the stream as it confronts a ferry connecting two ends of a highway at two opposite points across the river (Fig. 18-1). It is clear that when crossing the river, the ferry must keep its course slightly upstream to compensate for the drift. Thus, while it covers the distance AB with respect to the water, it drifts downstream the distance BC. Applying the Pythagorean theorem to the right triangle ABC, we can write

$$(AB)^2 = (BC)^2 + (AC)^2.$$

Now (merely in order to arrange things in a form a little more convenient for our purpose), let us multiply and divide BC by AB:

$$BC = \frac{BC \times AB}{AB} = AB \times \frac{BC}{AB}.$$

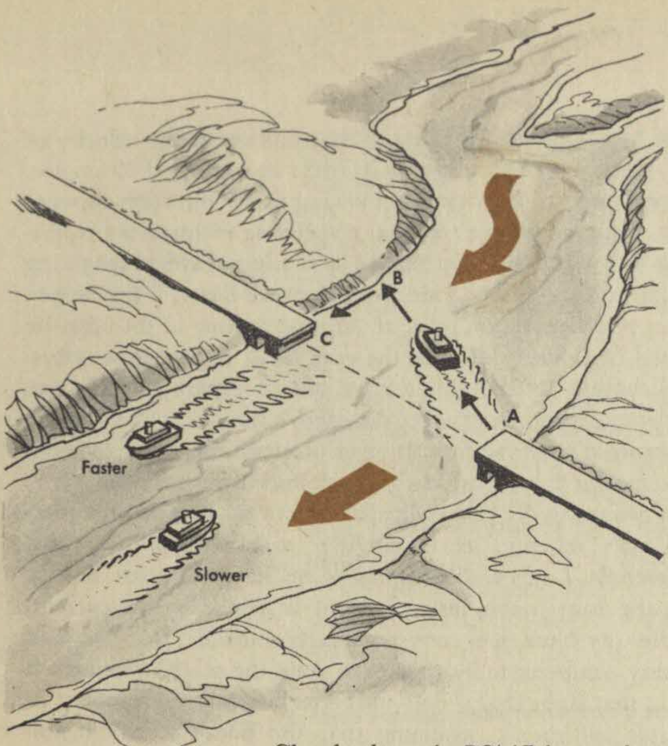


FIG. 18-1 How to compensate for drift in steering a boat directly across a moving stream.

Clearly the ratio BC/AB is equal to the ratio of the velocity of the river v_{R} to the velocity of the boat with respect to the moving water v_{B}. Thus we have

$$BC = AB \times \frac{v_{\mathrm{R}}}{v_{\mathrm{B}}}$$

and the original Pythagorean equation becomes

$$(AB)^2 = (AB)^2 \times \frac{v_{\mathrm{R}}^2}{v_{\mathrm{B}}^2} + (AC)^2$$

$$(AB)^2\left[1 - \left(\frac{v_{\mathrm{R}}}{v_{\mathrm{B}}}\right)^2\right] = (AC)^2.$$

Since we have assumed the ratio of the velocities is 1/10, we have

$$AB = \frac{AC}{\sqrt{1 - 0.01}} = 1.005\,AC.$$

Since this result applies equally well to both crossings, the distance, with respect to the water to be covered by our ferry on a round trip across the river, will also be 0.5 percent longer and so will be the time needed for the trip. Thus we find that moving across the stream also introduces a delay, but this delay is only half as much as the delay connected with sailing up and downstream.

Now substitute the "ether wind" for the river, and propagating light waves for the boats, and you will have the principle of Michelson's experiment. The details are shown in Fig. 18-2. A light beam from a source S falls on the glass plate P, which is covered with a thin semi-

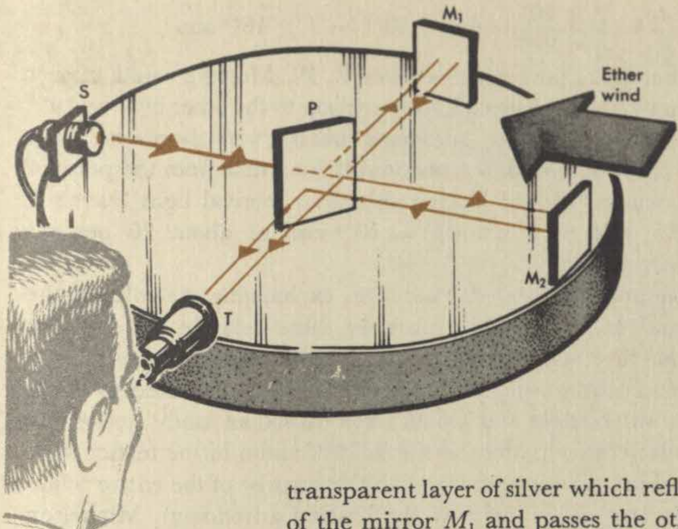


FIG. 18-2 Michelson's apparatus for trying to measure the speed of the earth's motion through the ether. Instruments were mounted on a heavy stone plate floating on mercury to avoid vibration and warping when the apparatus was rotated.

transparent layer of silver which reflects half of the beam in the direction of the mirror M_1 and passes the other half through in the direction of mirror M_2. Being reflected by these mirrors, the beams return back to the plate P; half of the first beam penetrates the thin silver coating on P and continues on to the telescope and the eye at T, and half of the second beam is reflected into the telescope by the silver layer. Thus the two beams entering the telescope will have the same intensity and will produce a clear, definite interference pattern in the field of view of the telescope.

However, in the presence of "ether wind," the situation was expected to be rather different. When the apparatus was placed in such a way that the line PM_2 coincided with the direction of the wind or its projection on a horizontal plane, the light waves traveling in this direction would be in the position of a boat sailing up and downstream, whereas the light waves traveling along the line PM_1 would correspond to a ferry moving to and fro across the river. Because of the difference in time delays in the two cases, the light beams would not arrive at the telescope simultaneously, and the difference in their times of arrival would result in a shift in the interference pattern. The ratio of the orbital velocity of the earth, 30 km/sec, to the velocity of light, 300,000 km/sec, is considerably smaller than in the nautical example discussed above. Using the same method of calculation, we find that in this case the two light beams should arrive at the telescope with a relative delay of only 5×10^{-9} of the total travel time.

Actually, Michelson's apparatus was a little more complicated than Fig. 18-2 indicates. Auxiliary mirrors reflected the beam back and forth several times between P and M_1, and between P and M_2, so that the total travel of the beams was equivalent to more than 10 m downstream and 10 m back, or 20 m across the stream and back. So the total travel time for each beam was

$$\frac{d}{v} = \frac{2 \times 10^3}{3 \times 10^{10}} = 0.7 \times 10^{-7} = 7 \times 10^{-8} \text{ sec.}$$

Therefore, Michelson (and his co-worker E. W. Morley) could expect that one beam would be delayed in comparison to the other by $7 \times 10^{-8} \times 5 \times 10^{-9} = 35 \times 10^{-17}$ sec. Although this is a very short time from the everyday point of view, it is a reasonably long time from the point of view of light waves. Indeed, during this time interval light travels a distance of $35 \times 10^{-17} \times 3 \times 10^{10} = 10^{-5}$ cm, or about 20 percent of a visible wavelength.

Turning the apparatus by 90° and thus exchanging the roles of the mirrors M_1 and M_2, we would expect the same delay in the opposite direction. Thus the total difference between the two light beams in the first and in the second positions of the apparatus was expected to be 40 percent of the wavelength and should have caused an easily noticeable shift in the interference pattern in the field of vision in the telescope.

However, to his great surprise, and to the surprise of the entire scientific world (at least in the fields of physics and astronomy), Michelson failed to notice any change at all! How could it be? Some scientists suggested that there might be some drag of the ether caused by the moving bulk of the earth so that the resulting velocity of the "ether wind" near the ground is considerably reduced. A British physicist, G. F. Fitzgerald (1851–1901), tried to interpret the negative result of Michelson's experiment by postulating that all material bodies moving through the ether shrink in the direction of their motion by an amount dependent on their velocity. H. A. Lorentz followed up Fitzgerald's interpretation by explaining this hypothetical contraction by a change of electric and magnetic forces between atoms moving through the ether. The effect of this Fitzgerald-Lorentz contraction would be to transform the round table on which Michelson's mirrors were mounted into an ellipse with the shorter axis in the direction of the earth's motion. This would reduce the distance to be traveled by the light beam propagating in the "up and downwind" direction and would enable that light beam to arrive at the telescope simultaneously with the beam that was traveling across the wind. Other attempts were made to relate the failure of the Michelson-Morley experiment to a relative motion between the apparatus and a "sea of ether," but they did not lead to positive results.

Michelson's failure to detect the motion of the earth through the ether had the same roots as the failure of contemporary physical theories to formulate the mechanical properties of this hypothetical medium. It was illogical to ascribe to the hypothetical ether the properties of ordinary matter, such as, for example, elasticity or compressibility, since in doing so we would also have to assume that ether possesses some kind of granular structure formed by "subatoms." But if, on the other hand, we consider the ether to be an absolutely homogeneous substance without any internal structure, there is no logical possibility of talking about the

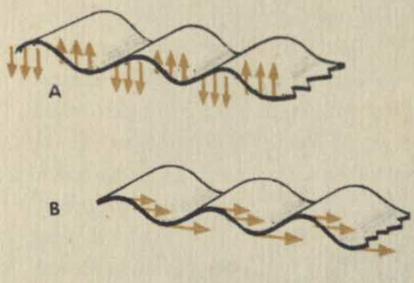


FIG. 18-3 (A) A wave propagating through a stationary elastic ribbon; (B) a moving wave-shaped rigid ribbon.

motion of that ether or the motion of objects with respect to it. In fact, when we watch a rotating disk, we notice that it rotates by observing the motion of minor marks on its surface, such as scratches or dents. If the surface of the disk is perfectly smooth, with no marks that would catch our eye, we shall not be able to tell just by looking at it whether it is moving or not. But, of course, we can touch it with our fingertip and immediately feel whether its surface is at rest or slides under our finger. And if the disk rotates fast enough, we shall feel the warmth produced by the friction between the skin of our finger and the moving surface of the disk. But the phenomenon of friction, which informs us about the state of motion of the disc, is again a purely molecular phenomenon and would be absent in an "absolutely homogenous" substance.

If we give a little thought to the problem, we can easily persuade ourselves that *it is meaningless to talk about the motion of a continuous medium or motion with respect to it unless this medium can be considered to be formed by individual discrete particles.* If we had a long ribbon made of an absolutely continuous material and observed a wave propagating along it (Fig. 18-3), it would be meaningless to ask whether it is (A) a regular elastic wave propagating along a stationary ribbon or (B) a rigid ribbon cut out of a sheet shaped like corrugated iron that is moving bodily from left to right.

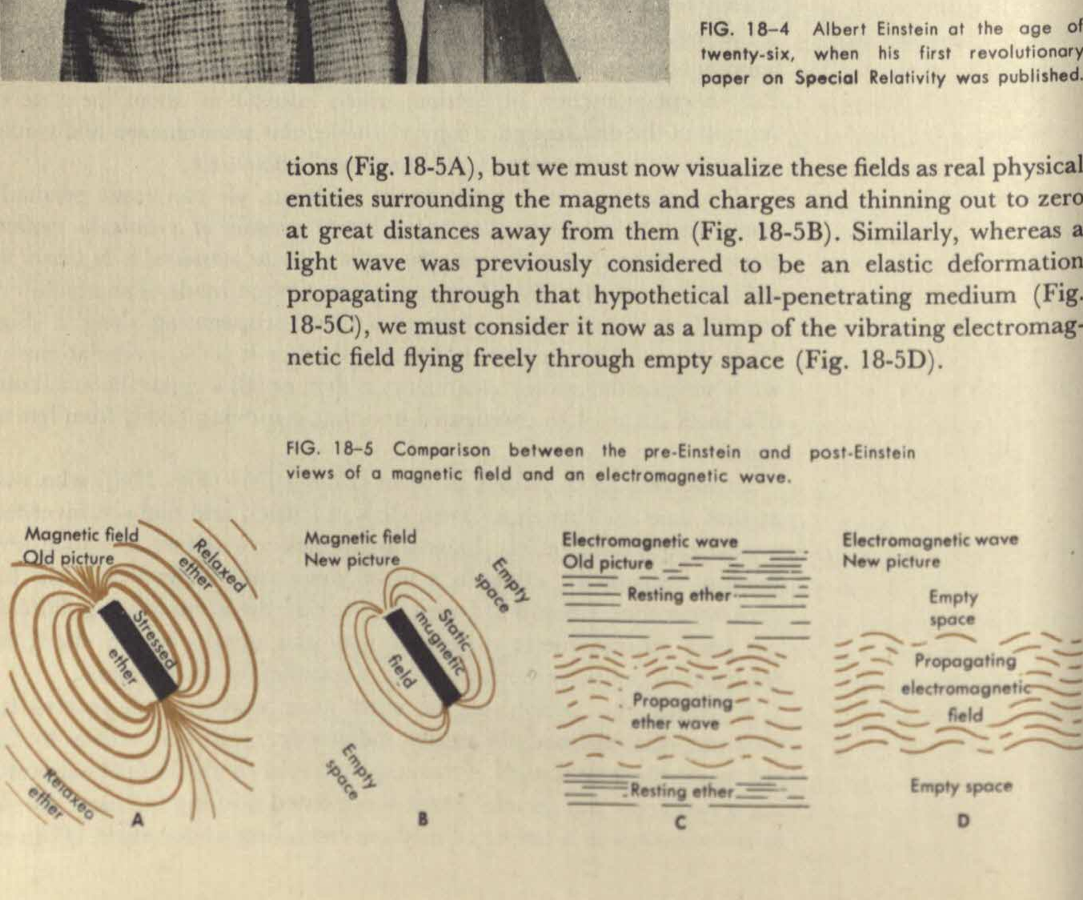
In the year 1905, Albert Einstein (1879–1955) (Fig. 18-4), who was at that time working as a patent clerk in Zurich and had just invented a new type of oil pump, looked at Michelson's failure to notice any motion through the ether in a much more radical way than did his contemporaries. Instead of trying to patch up the accumulating difficulties and contradictions connected with the notion of the ether, he returned to the pre-etherian idea of a completely empty space.

Rejecting the notion of the ether as a universal all-penetrating medium, Einstein had to ascribe independent physical reality to the electromagnetic field itself. According to the old views, the field surrounding a magnet or an electric charge represented nothing but local stresses or deformations in a universal medium extending uniformly in all direc-

FIG. 18–4 Albert Einstein at the age of twenty-six, when his first revolutionary paper on Special Relativity was published.

tions (Fig. 18-5A), but we must now visualize these fields as real physical entities surrounding the magnets and charges and thinning out to zero at great distances away from them (Fig. 18-5B). Similarly, whereas a light wave was previously considered to be an elastic deformation propagating through that hypothetical all-penetrating medium (Fig. 18-5C), we must consider it now as a lump of the vibrating electromagnetic field flying freely through empty space (Fig. 18-5D).

FIG. 18–5 Comparison between the pre-Einstein and post-Einstein views of a magnetic field and an electromagnetic wave.



18-3 Relativistic Mechanics

Along with the exit of the ether from the stage of physics, out went also the notion of "absolute motion" through space, which was always associated, though often subconsciously, with the idea of motion with respect to an ether. If there is no ether filling all of space and serving as a universal reference system for the motion of material bodies, we can speak only about the motion of one material body relative to another material body, and ***the basic laws of physics should be the same no matter in what system of reference we are studying them.*** Because of this basic postulate stating that there is no such thing as *absolute motion* and that only a *relative motion* of one object with respect to another has a physical meaning, Einstein's theory is commonly known as the *theory of relativity*.

It follows from the above postulate that it should be impossible to detect the motion of one system of reference with respect to another by performing some physical experiment in each of them and then comparing the results. Thus the grandfather's clock in the captain's cabin of an ocean liner speeding toward New York across the smooth, blue waters of the Atlantic with no storm or choppy sea breaking the uniformity of motion will operate just as well as it would standing in the sitting room of the captain's home. And the passengers playing Ping Pong or billiards on this ship will not be able to tell whether their ship is lying quietly in the Southampton docks or sailing across the Atlantic. Michelson's experiment had shown that this is also true for light phenomena. A physicist repeating Michelson's experiment in an inside cabin of the ship will not know whether the ship is moving or resting (relative to the dry land) unless he goes up on deck and sees the gray buildings of dock installations or the limitless expanses of the ocean.

Einstein, instead of trying to explain away the negative result of the Michelson-Morley experiment, merely accepted it at its full face value: ***the speed of light (in empty space) is the same for all observers, no matter what their velocity, or what the velocity of the source.***

This second postulate is not always too easy to accept. If a rifleman sits in a fast-moving jeep and shoots in a forward direction, the velocity of the bullets with respect to the ground will be the sum of the muzzle velocity and the velocity of the jeep (Fig. 18-6A), but bullets shot backward will fly correspondingly slower (Fig. 18-6B). If we consider light as some kind of vibrating bullets emitted by light sources, we would expect that the velocity of light emitted by an approaching source would be higher than that emitted by a receding source. Abundant astronomical evidence based on the observation of binary stars proves, however, beyond any doubt that this is not the case. A binary star (Fig. 18-6C) is a system of two giant suns rotating around their common center of gravity and is a rather common object in the sky (in fact, about half of all known stars are binaries). Because of its rotation around the common center, each of the stars is moving toward us during half of its rotation period and away from us during the other half.

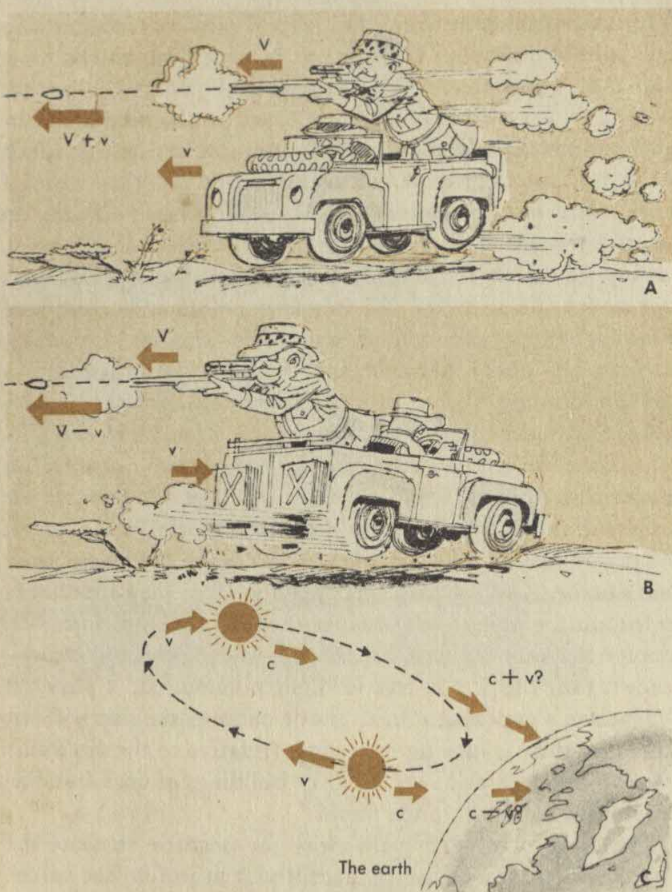


FIG. 18–6 The velocity of a bullet is affected by the motion of its source. Is this also true in the case of light?

If the velocity of light were affected by the motion of the source, the light from the approaching star would reach us sooner than the light from the receding one, and the difference in the arrival of the two light signals would be quite large. Assuming, for example, that the orbital velocities of the two stars are the same as the orbital velocity of the earth, i.e., 30 km/sec (and they are often larger than that), we find that the light would be accelerated or retarded by 0.01 percent, depending on whether it comes from the approaching or from the receding component of a binary star. Over a distance of 100 light-years, which is not uncommon for the observed binaries, this seemingly small difference in velocity could result in a week's difference between the arrival of light from these two stars to the earth, and this difference would be reversed every half revolution period. Thus an astronomer observing a binary star would find himself in the position of a sports fan watching a prize fight on a TV

screen which, because of some trouble in transmission during the third round, shows the champion and the challenger with a few minutes difference in phase. Our fan would see the champion *already* resting in his corner while the challenger is *still* shadow boxing in the middle of the arena, and a minute later the champion would go out for the kill while the challenger is *still* being readied by his seconds. In the middle of the fourth round, the fight would seem to be normal; but toward the end of it, things would change in the opposite direction, and the surprised sports fan would see the champion deliver his K.O. blow after the challenger was already counted out. Since nothing of this kind has ever been seen by astronomers observing the motion of binary stars, we must conclude that the velocity of light is not affected by the motion of its source.

And, according to Einstein's first postulate, it does not matter whether we want to consider the star to be approaching us or ourselves to be approaching the star; the relative velocity is all that matters. This, then, means that if we add the velocity of light to any other velocity, we get again the same original velocity of light! This contradicts common sense! Well, said Einstein, if there is a scientifically established paradox you cannot get rid of, all you can do is to rationalize it. And as for the common sense . . . well, the same common sense was once objecting to the idea that the earth is round. If the commonsense idea concerning the addition of two velocities does not apply to the velocity of light and the velocity of its source, it must be generally wrong and its use in everyday life may be justified only by the fact that all the velocities we encounter in ordinary life are much smaller than that of light. Thus, cutting another Gordian knot, Einstein introduced a new and at first sight very strange law governing the addition of two velocities. If v is the velocity of the jeep and V the muzzle velocity of the bullet shot in the forward direction by the rifleman in the jeep, the velocity of the bullet with respect to the ground will be, not $V + v$, but

$$\frac{V+v}{1+\frac{V\times v}{c^2}}$$

where c is the velocity of light. If both velocities, V and v, are small compared with the velocity of light, the second term in the denominator is practically zero, and the old "commonsense" formula holds. But if either V or v, or both, approaches the velocity of light c, the situation will be quite different.

Suppose that the velocity of the jeep is 75 percent of the speed of light and that the muzzle velocity of the rifleman's bullet is the same. According to common sense, the velocity of the bullet with respect to the ground should be 50 percent above the velocity of light. However,

putting $V = 0.75c$ and $v = 0.75c$ into the above formula, we get only $0.96c$, so that the velocity of the bullet with respect to the ground remains less than the speed of light. The reader can easily verify this fact; no matter how close the two velocities to be added are to the velocity of light, the resulting velocity will never exceed it. In the limiting case, if we make $v = c$, we obtain

$$\frac{V + c}{1 + \frac{V \times c}{c^2}} = \frac{V + c}{1 + \frac{V}{c}} = \frac{c(V + c)}{(c + V)} = c.$$

This equation puts a quantitative foundation under the idea that the velocity of the source does not add anything to the velocity of light emitted by it. Fantastic as it may look at first sight, Einstein's law for the addition of two velocities is correct and has been confirmed by direct experiments. It does not agree with commonsense conclusions, but we should not forget that commonsense conclusions are based on our everyday experience, and neither a jeep traveling with a speed close to that of light nor rifles shooting bullets at that speed can be considered as an "everyday experience"! Thus Einstein's theory of relativity leads us to the conclusion that ***it is impossible to exceed the velocity of light by adding two (or more) velocities, no matter how close each of these velocities is to that of light.*** The velocity of light, therefore, assumes the role of some kind of *universal speed limit* which cannot be exceeded no matter what we do.

It is plain to see that if the speed of light (or of any electromagnetic radiation) is to be the same for all observers, no matter what their velocity relative to the source, then a law such as Einstein's is necessary for the addition of velocities. If this were all there was to it, the idea of "relativity" would have had but little effect. However, it is not possible to put forth such a law for the addition of velocities and leave the rest of physics unchanged. Velocity involves a combination of distance and time; and for any observer, velocity (as *he* measures it) must be equal to distance (as *he* measures it) divided by time (as *he* measures it).

If this last statement is also to be true, a number of other assumptions *necessarily* follow. They all can be demonstrated, and with only straightforward simple algebra; though, because a great deal of algebra is required, we shall not try to lead the reader through its mazes.

Imagine an observer O who considers himself to be stationary, as he has every right to do. He sees, passing by in the air, a spaceship-laboratory traveling with a velocity v relative to the observer. Through its large windows, O sees an experimenter carrying out some simple laboratory procedures in mechanics. He is equipped with standard masses plainly marked "1 kg," "100 gm," etc.; meter sticks; and a fine clock on the wall. The name of the manufacturer is visible on all this equip-

ment, and O has no reason to believe they are not proper standards, as marked.

Observer O, however, with his own ingenious equipment (and making proper allowance for the speed of light as it moves to him from the traveler) discovers the following:

O's measurements show that all the traveler's meter sticks are too short: instead of 1 meter, they are only $\sqrt{1 - (v^2/c^2)}$ meter long.

O finds the traveler's clock is running slow: a period of 1 minute by the traveler's clock is actually $1/\sqrt{1 - (v^2/c^2)}$ minute by the stationary clock in O's laboratory.

After watching some experiments, O comes to the further conclusion that the traveling kilograms are more massive than they are labeled: their actual masses are $1/\sqrt{1 - (v^2/c^2)}$ kg.

In all fairness, it should be noted that the spaceship experimenter has been closely observing what goes on in O's lab, and comes to exactly the same conclusions: that O's meter sticks are too short by the same factor, that O's clock is running slow, and that O's kilograms are really more massive than they are marked to be.

That each of these experimenters came to exactly the same conclusion about the other is just what we should have expected; the relative velocity between the two is what makes the difference, and this is the same for both. (Remember that O, presumably on the earth, who chose to consider himself stationary, was actually spinning several hundred miles per hour with the daily rotation of the earth, about 60,000 mi/hr in his orbiting around the sun, and at about 600,000 mi/hr in the solar system's revolution about the center of our Milky Way Galaxy!)

This little fable is not one that could ever be carried out in actuality. But many pieces of it have been observed to be true, and there is probably nothing in modern science more firmly established than Einstein's Special Theory of Relativity. In later sections on nuclear physics, we shall see much of the evidence.

What is true for meter sticks and clocks and mass standards is of course true for all lengths and time intervals and masses aboard any speeding traveler. At ordinary everyday speeds these relativistic changes are too small to be measurable. As an example, let us calculate a few things about an artificial satellite circling the earth at a speed of 11 km/sec = 1.1×10^6 cm/sec. Its rest mass m_0 (i.e., its mass as measured when stationary with respect to the measurer) is, say, exactly 1000 kg, and its rest length l_0 is exactly 5 m. Relativity shows us that a stationary earthly observer would measure its mass and length as

$$m = \frac{m_0}{\sqrt{1 - (v^2/c^2)}}$$

$$l = l_0 \times \sqrt{1 - (v^2/c^2)}.$$

Thus, while the satellite is circling, we would measure its mass (by any means we can imagine) to be

$$m = \frac{10^3}{\sqrt{1-(1.1\times 10^6/3\times 10^{10})^2}} = \frac{10^3}{\sqrt{1-1.4\times 10^{-9}}}.$$

This expression would be a very inconvenient one to evaluate by regular longhand means, so let us resort to an algebraic approximation that is only applicable to special cases:

$$\sqrt{1-\alpha} \approx 1 - \tfrac{1}{2}\alpha$$

ONLY IF α IS VERY MUCH SMALLER THAN 1.

Certainly 1.4×10^{-9} *is* very much smaller than 1, and so we shall be almost exactly right to say

$$m = \frac{10^3}{1-0.7\times 10^{-9}} = \frac{10^3}{1-7\times 10^{-10}}.$$

The division of this fraction would likewise be a clumsy piece of arithmetic, which we can again get around by an algebraic trick:

$$\frac{1}{1-\alpha} \approx 1 + \alpha$$

ONLY IF α IS VERY MUCH SMALLER THAN 1.

This gives us that

$$\begin{aligned} m &= 10^3 \times (1 + 7\times 10^{-10}) \\ &= 1000 + 7\times 10^{-7}\ \text{kg}. \end{aligned}$$

The *increase* in mass is thus 7×10^{-7} kg $= 7 \times 10^{-4}$ gm $= 0.7$ mg, certainly not measurable on a 1-ton object.

The length of the satellite, as we would measure it, will prove out in the same way

$$\begin{aligned} l &= l_0\sqrt{1-(v^2/c^2)} \\ &= 5 \times (1 - 7\times 10^{-10}) \\ &= 5 - 3.5\times 10^{-9}\ \text{m}. \end{aligned}$$

The shortening would thus turn out to be only 3.5×10^{-9} m $= 3.5 \times 10^{-7}$ cm $= 35$ Å, or about 1 percent of the wavelength of ultraviolet light!

If there were an accurate clock in this satellite, accurately set before it took off, how far off would this clock be when the satellite returned safely to earth after circling for, say, 10 days? From our point of view here on earth, the traveling clock will be running slow by a factor of $\sqrt{1-(v^2/c^2)} = 1 - 7 \times 10^{-10}$. In 10 days there are $10 \times 24 \times 60 \times 60 = 8.64 \times 10^5$ sec. The time interval recorded by the satellite clock would then be

$$t = 8.64 \times 10^5(1 - 7 \times 10^{-10})$$
$$= 8.64 \times 10^5 - 6 \times 10^{-4} \text{ sec.}$$

That is, it would be slow (according to our stay-at-home standards) by only 6×10^{-4} sec.

18-4 Space-Time Transformation

Einstein's new laws clearly contradict the classical (commonsense) ideas concerning space and time, so that in accepting these new laws as experimental fact, we are forced to introduce radical changes in our old notions. In his *Principia*, the great Newton wrote:

> ***I. Absolute, true, and mathematical time, of itself, and from its own nature, flows equably without relation to anything external.***
>
> ***II. Absolute space, in its own nature, without relation to anything external, remains always similar and immovable.***

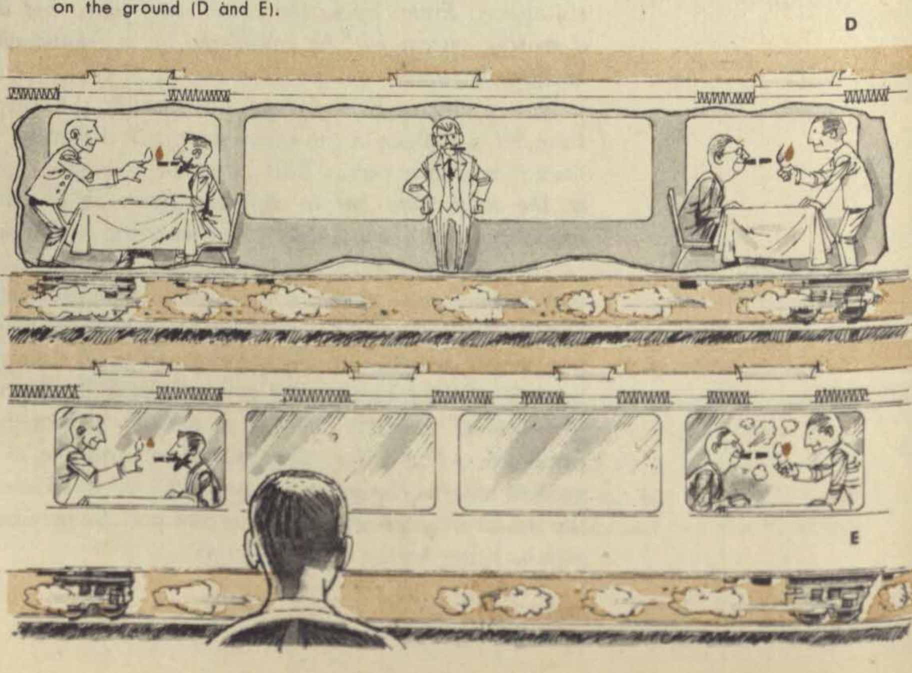
According to Einstein's views, however, space and time are more intimately connected with one another than it was supposed before, and, within certain limits, the notion of space may be interchanged with the notion of time, and vice versa. To make this statement more clear, let us consider a railroad passenger having his meal in the dining car. The waiter serving him will know that the passenger ate his soup, steak, and dessert in the same place, i.e., at the same table in the car. But, from the point of view of a person on the ground, the same passenger consumed the three courses at points along the track separated by many miles (Fig. 18-7A, B, and C). Thus we can make the following trivial statement: ***Events occurring in the same place but at different times in a moving system will be considered by a ground observer as occurring at different places.***

Now, following Einstein's idea concerning the reciprocity of space and time, let us replace in the above statement the word "place" by the word "time" and vice versa. The statement will now read: ***Events occurring at the same time but in different places in a moving system will be considered by a ground observer as occurring at different times.***

This statement is far from being trivial and means that if, for example, two passengers at the far ends of the diner had their after-dinner cigars lighted simultaneously from the point of view of the dining-car steward, the person standing on the ground will insist that the two cigars were lighted at different times (Fig. 18-7D and E). Since, according to the principle of relativity, neither of the two reference systems should be preferred to the other (the train moves relative to the ground or the ground moves relative to the train), we do not have any reason to take the steward's impression as being true and the ground observer's impression as being wrong, or vice versa.

Why then do we consider the transformation of the time interval (between the soup and the dessert) into the space interval (the distance along the track) as quite natural, and the transformation of the space interval (the distance between the two passengers having their cigars lit) into the time interval (between these two events as observed from the track) as paradoxical and very unusual? The reason lies in the fact that in our everyday life we are accustomed to velocities that lie in the lowest brackets of all the physically possible velocities extending from zero to the velocity of light. A race horse can hardly do better than about one millionth of a percent of this upper limit of all possible velocities, and a modern supersonic jet plane makes, at best, 0.0003 percent of it. In comparing space and time intervals, i.e., distance and durations, it is rational to choose the units in which the limiting velocity of light is taken to be 1. Thus, if we choose a "year" as the unit of duration, the corresponding unit of length will be a light-year, or 10,000,000,000,000 km, and if we choose a "kilometer" as the unit of length, the unit of time will be 0.000003 sec, which is the time interval necessary for light to cover the distance of 1 km. We notice that whenever we choose one unit in a "reasonable" way (a "year" or a "kilometer"), the other unit comes out either too large (a light-year) or too short (3 microseconds) from the point of view of our everyday experience. So, in the case of the

FIG. 18–7 Events that occur at the same place on a moving railway car will occur at different locations as observed from the ground (A, B, and C). Events that occur simultaneously to an observer on the moving car will appear to take place at different times to an observer on the ground (D and E).



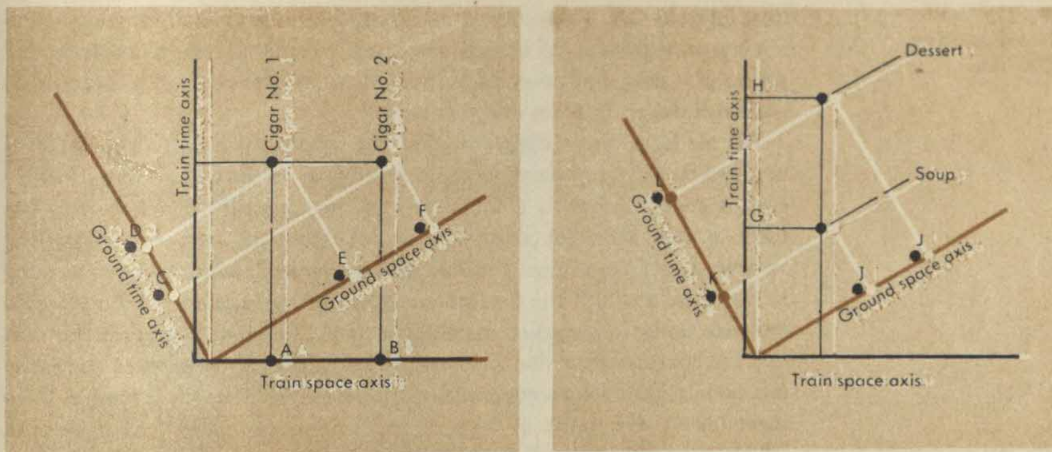


FIG. 18-8 The difference in space and time separations between events, as interpreted by observers moving relative to one another.

passenger eating his dinner on the train, a half-hour interval between the soup and the dessert could result in 200,000,000 mi of distance along the track (time $\times$ c) if the train were moving at a speed close to that of light, and we are not surprised that the actual difference is only 20 or 30 mi. On the other hand, the distance of, let us say, 30 m between two passengers lighting their cigars at opposite ends of the railroad car translates into a time interval of only one hundred-millionth of a second (distance $\div$ c), and there is no wonder that this is not apparent to our senses.

The transformation of time intervals into space intervals and vice versa can be given a simple geometrical interpretation, as was first done by the German mathematician H. Minkowski, one of the early followers of Einstein's revolutionary ideas. Minkowski proposed that time or duration be considered as the fourth dimension supplementing the three spatial dimensions and that the transformation from one system of reference to another be considered as a rotation of coordinate systems in this four-dimensional space. His basic idea can be understood by considering the diagrams shown in Fig. 18-8. From the point of view of the observer on the train, the time interval between the two cigar lightings is zero, and they are separated in space by the length of the dining car AB. To the observer on the ground, there will be a time interval CD between the two lightings, and they will also appear to be separated by a shorter distance EF. To the train observer, both the soup and the dessert will be eaten in the same place and will be separated by the

time interval GH. The observer watching from the ground, however, will not agree. He will, of course, say they were eaten in quite different places (IJ) and will contend that the time interval between them (KL) is shorter than the train observer says.

We see from these diagrams that the appearance of a time interval between two events which were simultaneous in the first system of reference is connected with a shortening of the apparent distance between them as seen from the second system of reference, and vice versa; the appearance of a space interval between two events which were occurring in the same place of the first system shortens the apparent time interval between them as observed from the second. The first fact gives the correct interpretation of the apparent Fitzgerald's contraction of the moving bodies, and the second makes the time in a moving system flow slower from the point of view of the second system. Of course, both effects are relative, and each of the two observers moving with respect to one another will measure the other one to be somewhat flattened in the direction of his motion and will consider the other fellow's watch to be slow.

18-5 Mr. Tompkins

Because these effects become appreciable only when the velocities involved are close to that of light, we do not notice them at all in our everyday snail's-pace life. But we can imagine a fictitious situation which would arise if the velocity of light were much smaller and closer to our everyday experience. This is what happened to Mr. Cyril George Henry Tompkins, who after listening to a popular lecture on the theory of relativity, was transferred, in his dream, to a fantastic city in which the velocity of light was only 20 miles per hour and served as the natural speed limit for its inhabitants.*

> At first, when he found himself on the street of this relativistic city, nothing unusual seemed to be happening around him; even a policeman standing on the opposite corner looked as policemen usually do. The hands of the big clock on the tower down the street were pointing almost to noon and the streets were nearly empty. A single cyclist was coming slowly down the street and, as he approached, Mr. Tompkins' eyes opened wide with astonishment. For the bicycle and the young man on it were unbelievably flattened in the direction of the motion, as if seen through a cylindrical lens (Fig. 18-9). The clock on the tower struck twelve, and the cyclist, evidently in a hurry, stepped harder on the pedals. Mr. Tompkins did not notice that he gained much in speed, but, as the result of his effort, he flattened still more and went down the street looking exactly like

* From G. Gamow, *Mr. Tompkins in Wonderland* (Cambridge: Cambridge University Press, 1939), by permission of the publishers. Certain slight modifications have been made in order to place Mr. Tompkins' dream city in America rather than in England.

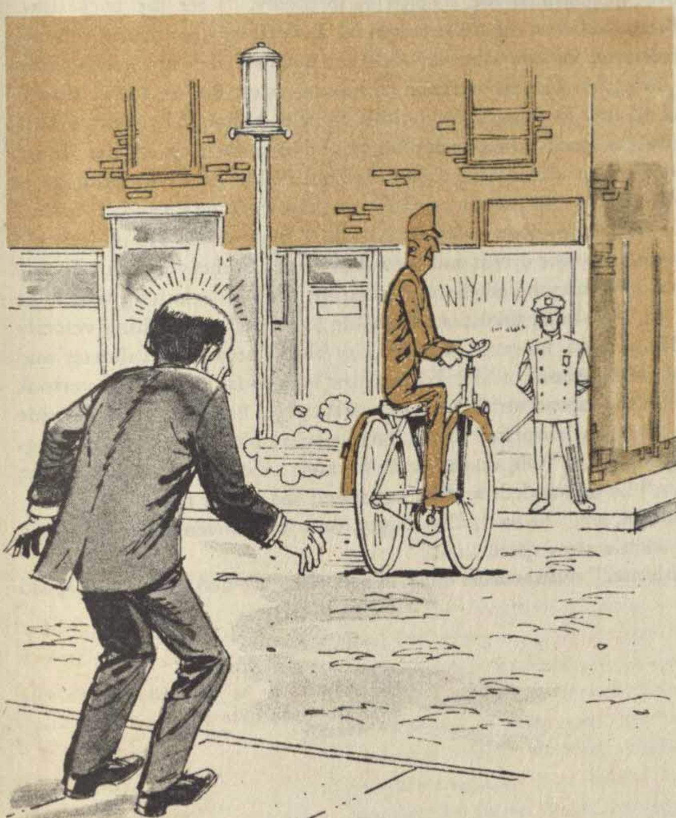


FIG. 18-9 The bicyclist seemed to be unbelievably flattened.

a picture cut out of cardboard. Mr. Tompkins felt very proud because he could understand what was happening to the cyclist—it was simply the contraction of moving bodies. "Evidently nature's speed limit is quite low here," thought Mr. Tompkins, "that is why the policeman on the corner looks so lazy; he need not watch for speeders." In fact, a taxi moving along the street at the moment and making all the noise in the world could not do much better than the cyclist, and was just crawling along. Mr. Tompkins decided to overtake the cyclist, who looked a good sort of fellow, and ask him all about it. Making sure that the policeman was looking the other way, he borrowed somebody's bicycle standing near the curb and sped down the street. He expected that he would be immediately flattened, and was very happy about it as his increasing figure had lately caused him some anxiety. To his great surprise, however, nothing happened to him or to his cycle. On the other hand, the picture around him completely changed. The streets grew shorter, the windows of the shops began to look like narrow slits, and the policeman on the corner became the thinnest man he had ever seen (Fig. 18-10).

"By Jove!" exclaimed Mr. Tompkins excitedly, "I see the trick now. This is where the word relativity comes in. Everything that moves relative to me gets shorter for me, whoever works the pedals!" He was a good cyclist and was doing his best to overtake the young man. But he found that it was not at all easy to get up speed on this bicycle. Although he was working on the pedals as hard as he possibly could, the increase in speed was almost negligible. His legs already began to ache, but still he could not manage to pass a lamp post on the corner much faster than when he started. It looked as if all his efforts to move faster were leading to no result. He understood now very well why the cyclist and the cab he had just met could not do any better, and he remembered the words in the book on relativity which he had read. It was stated that it is impossible to surpass the limiting velocity of light. He noticed, however, that the city blocks became still shorter and the cyclist riding ahead of him did not now look so far away. He overtook the cyclist at the second turning and, when they had been riding side by side for a moment, was surprised to see that he was quite a normal, sporting-looking young man. "Oh, that must be because we do not move relative to each other," he concluded; and he addressed the young man.

"Excuse me, sir!" he said, "Don't you find it inconvenient to live in a city with such a slow speed limit?"

"Speed limit?" returned the other in surprise, "we don't have any speed

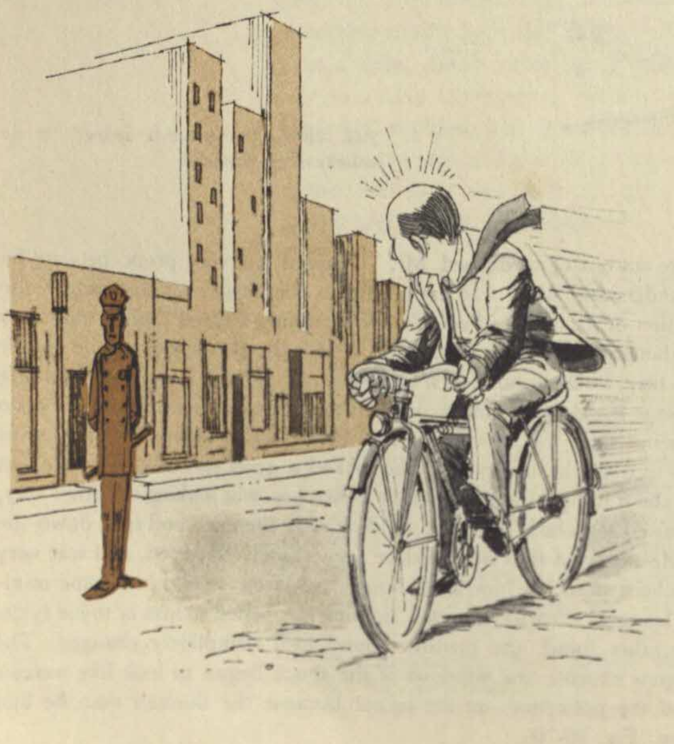


FIG. 18-10 As he speeded up, the city blocks became quite short.

limit here. I can get anywhere as fast as I wish, or at least I could if I had a motorcyle instead of this good-for-nothing old bike!"

"But you were moving very slowly when you passed me a moment ago," said Mr. Tompkins. "I noticed you particularly."

"Oh you did, did you?" said the young man, evidently offended. "I suppose you haven't noticed that since you first addressed me we have passed five blocks. Isn't that fast enough for you?"

"But the blocks became so short," argued Mr. Tompkins.

"What difference does it make, anyway, whether we move faster or whether the blocks become shorter? I have to go ten blocks to get to the post office, and if I step harder on the pedals the blocks become shorter and I get there quicker. In fact, here we are," said the young man, getting off his bike.

Mr. Tompkins looked at the post office clock, which showed half-past twelve. "Well!" he remarked triumphantly, "it took you half an hour to go this ten blocks, anyhow—when I saw you first it was exactly noon!"

"And did you *notice* this half hour?" asked his companion. Mr. Tompkins had to agree that it had really seemed to him only a few minutes. Moreover, looking at his wrist watch he saw that it was showing only five minutes past twelve. "Oh!" he said, "is this post office clock fast?" "Of course it is, or your watch is too slow, just because you have been going too fast. What's the matter with you anyway? Did you fall down from the moon?" and the young man went into the post office.

Continuing his journey down the street he finally saw the railway station. A gentleman obviously in his forties got out of the train and began to move towards the exit. He was met by a very old lady, who, to Mr. Tompkins' great surprise, addressed him as "dear Grandfather." This was too much for Mr. Tompkins. Under the excuse of helping with the luggage, he started a conversation.

"Excuse me, if I am intruding into your family affairs," said he, "but are you really the grandfather of this nice old lady? You see, I am a stranger here, and I never . . ." "Oh, I see," said the gentleman, smiling through his moustache. "I suppose you are taking me for the Wandering Jew or something. But the thing is really quite simple. My business requires me to travel quite a lot, and, as I spend most of my life in the train, I naturally grow old much more slowly than my relatives living in the city. I am so glad that I came back in time to see my dear little granddaughter still alive! But excuse me, please, I have to help her into the taxi," and he hurried away, leaving Mr. Tompkins alone again with his problems.

18-6 Time and the Space Traveler

In the Tompkins story, the relation between the young grandfather and his old granddaughter is, of course, grossly exaggerated, but the fact is that, according to Einstein's theory, such a difference in aging is really expected to occur in the case of relative motion. Thus, if sometime in the future a spaceship were to take off from the surface of the

earth to visit other planets of the solar system or maybe the planetary systems of other stars of the Milky Way, the pilot and the passengers would be relatively younger upon their return than the people of the same original age who had stayed on the earth. This difference might become quite conspicuous if the spaceship were accelerated to velocities close to the velocity of light. It is well known, however, that human organisms cannot stand strong accelerations and that pilots suffer black-out when their plane makes several g's ($g = 980$ cm/sec^2 being the normal acceleration of gravity on the surface of the earth). If we assume that the spaceship is traveling with the comfortable acceleration of 1 g, we find that it takes about a year to approach the velocity of light, when relativistic changes of time rate begin to play any role. (Since a year contains about 3×10^7 sec, the acceleration of 980 cm/sec^2 will raise the velocity to 98 per cent of the speed of light.) For accelerated space trips which last well beyond 1 yr, such as a trip to the nearby star Sirius, which is 8 light-years away, relativistic time changes begin to be quite appreciable. From the point of view of the inhabitants of the earth, the crew of a ship making a round trip to this star comes back just 16 yr after the departure; for the crew itself, these 16 yr will seem only as 9 yr. If, instead of to a nearby star, the spaceship travels with constant acceleration to the center of our own stellar system of the Milky Way and back, it will return 40,000 yr later by the earth's calendar, whereas by its own time reckoning, the trip will take only 30 yr.

Earlier in this chapter, however, it was said that *each* observer would think the clock of the other one was slow. Why, then, could we not equally well argue that when he landed, the traveler would find that the earth clocks were slow and the error would be the other way? To help resolve this argument, we can note that the two are *not* exactly equivalent. Earth has been virtually unaccelerated all the time; the satellite has undergone a takeoff acceleration, a turn-around acceleration, and a landing deceleration. There have been arguments over this question for many years, but now almost all relativists agree that the satellite clock *would* come back reading slow relative to the unaccelerated earth clock. The traveler to Sirius and back *would* return younger than his twin brother who stayed at home.

18-7 Energy-Mass Transformation

Perhaps the most important consequence of the Special Theory of Relativity is one that has not yet been mentioned. Einstein's famous equation

$$E = mc^2$$

has been widely publicized in many ways, and there are few people

who cannot recall at least having seen it somewhere. Its application has helped shape the modern world and will no doubt have an equally powerful effect on the future.

Our hypothetical observer O, in calculating from the data observed in the passing spaceship, will be unable to reconcile his observations with the laws of conservation of energy and conservation of momentum unless he adds another item to his list. This item is Einstein's $E = mc^2$, which basically says merely that energy has mass. Although the derivation in Einstein's original work referred only to kinetic energy, Einstein assumed that the relationship must also apply to energy of all kinds. Later experiments have proved him to be right. We now know that, not only does energy have mass, but that energy can be converted into mass and mass can be converted into energy.

In the equation, c^2 (the square of the speed of light) is a very large number, which means that a small amount of mass corresponds to a large amount of energy. In the CGS system ($c = 3 \times 10^{10}$ cm/sec), one gram of mass is interchangeable with 9×10^{20} ergs of energy. In the MKS system, one kilogram of mass is interconvertible with 9×10^{16} joules of energy.

Because of the large value of the proportionality constant c^2, the mass of a given amount of energy is usually very small. Thus the mass of light emitted per minute by a flashlight with a 10-watt bulb amounts to only 7×10^{-12} gm; the mass of the magnetic field surrounding an ordinary laboratory magnet is 10^{-15} gm, and the heat which has to be supplied to 1 kg of water to raise its temperature from the freezing point to the boiling point (10^5 cal) has a mass of about 4×10^{-9} gm.

On the other hand, when a certain amount of mass is transformed into energy, the situation is reversed (we now *multiply* mass by c^2), and we obtain a great deal of energy from a very small mass change. Thus, for example, the uranium core of the first atomic bomb lost only about 1 gm of its original mass in the course of its transformation into fission products. Being turned into energy, this gram of mass produced an effect equivalent to the explosion of 20,000 tons (20 kilotons) of TNT.

18-8 The Appearance of Moving Objects

It is not possible for any of us to actually grasp or have a clear mental picture of the speed of light, 3×10^{10} cm/sec or 186,000 mi/sec. The Tompkins story has the great advantage of reducing the speed of light to 20 mi/hr, which is something our experience qualifies us to deal with. But the story does also have a major defect. It was written in 1939, and it was not until 20 years later that any relativist had considered what we would actually *see* if some familiar object were to whiz by at a very

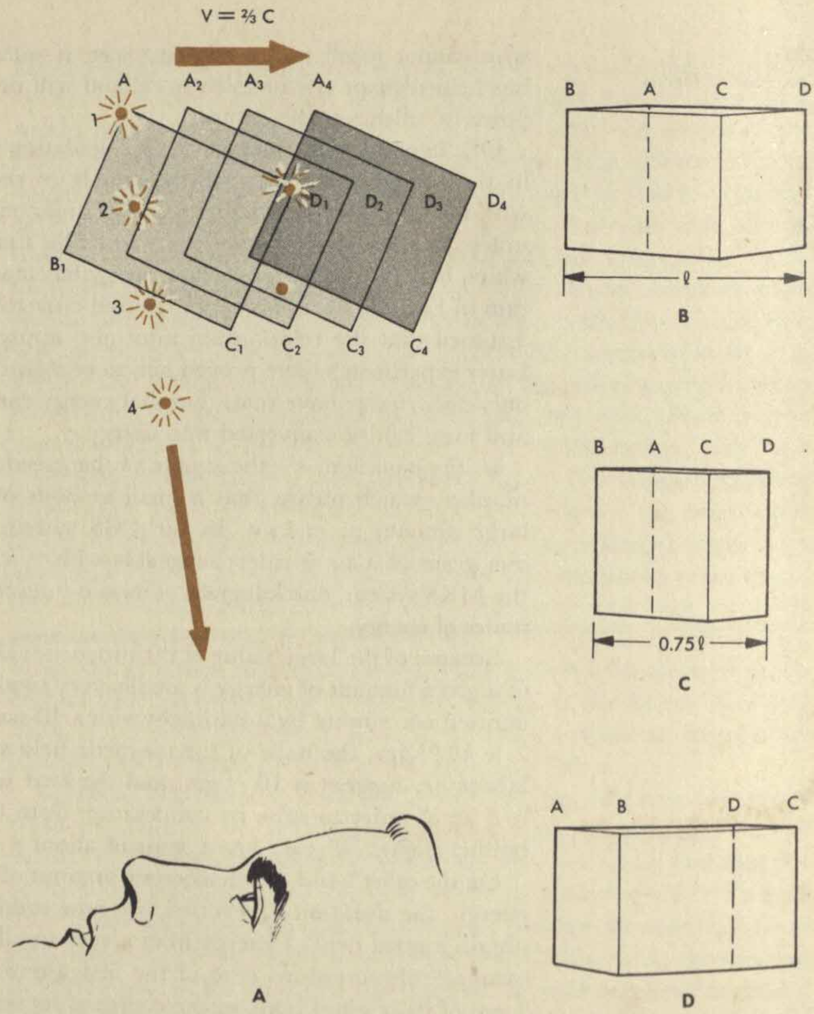


FIG. 18–11 Appearance of a moving block. The observer will not see the actual measured shortening; the block will instead appear to be rotated.

high speed. In 1959, Prof. J. L. Terrell published in the *Physical Review* an analysis of how such fast-moving objects would appear.*

Figure 18-11A shows the top view of a rectangular block *ABCD* moving from left to right at a speed $v = \frac{2}{3}c$, together with an observer. If the block were stationary, the observer would see it as shown in Fig. 18-11B. Now, if the block were moving from left to right at two-

---

* Also see V. L. Weisskopf, "The Visual Appearance of Rapidly Moving Objects," *Physics Today*, **13** (1960), 24.

thirds the speed of light, we know that the observer would conclude that all dimensions in the direction of motion would be reduced to 0.75 of their stationary values. We might therefore assume (as did all of the early relativists) that the moving block would appear as shown in Fig. 18-11C: it has merely been shortened in its direction of motion. This is the assumption on which the artist based Figs. 18-9 and 18-10. However, no matter what Mr. Tompkins might *figure* about the cyclist and the policeman, what he would actually *see* is something quite different.

Imagine a ray of light starting toward the observer from the back edge A when the block is in position 1. This ray could never reach the observer's eye if the block were stationary; but since it is moving at high speed, the results will be different. In some tiny fraction of a second, the light from A_1 heading toward the observer would have reached point 2 if the block were not in the way; during the same small time interval, the block would have indeed moved out of the way of the light to its second position on the drawing, leaving the ray unobstructed. After another small time interval, the light ray would have reached point 3 and the block (now in its third position) obviously no longer threatens to obscure the light, which is now quite free to proceed on toward the observer's eye.

Similarly, a ray from D_1 has a clear path to the observer from a stationary block. However, when the block is moving at high speed, the ray can never get there. In the same small interval of time, light from D_1 on its way to the observer would have to have reached the point marked by the black dot; by this time, though, the block would have moved to its second position, where it has clearly intercepted the ray and prevented it from ever reaching the observer. The observer of the moving block would thus not see edge D at all, and edge A would be clearly in view to the left of B.

A mathematical analysis (which we shall not inflict on the reader) shows that the moving block would not *appear* to be shortened but would instead look as though it had been rotated, as in Fig. 18-11D. Fundamentally, of course, the cause of this illusion is that light rays which arrive simultaneously at the observer's eye left the moving block at different times. The ray from A, for example, has farther to go than the ray from C and must therefore have started from the block earlier, when it had not yet moved so far to the right.

The calculations can be done in reverse too; if the observer (who sees something like Fig. 18-11D) were to make all the proper allowances for the motion of the block and the differing distances to its various edges, he would finally conclude that the light had emanated from a shortened block such as that in Fig. 18-11C. That is, if he figured back from his observations, he would in this way *measure* the block to be only 0.75 as long as its stationary length.

Thus Mr. Tompkins would actually *see* a cyclist and a policeman of perfectly normal proportions, who would merely appear to be considerably rotated from their actual positions. Figures 18-9 and 18-10 are really all right after all—except that they do not portray what Mr. Tompkins would *see*, but rather what he could *calculate* about what he observed. The astonished expression on his face is still quite justifiable!

Questions

(18-2)

1. Consider a river 2 miles wide that flows 3 mi/hr, and a motorboat that travels at exactly 24 mi/hr. (a) How long will it take the boat to cross the river and return? (b) How long will it take it to go 2 miles downstream, and then back?

2. A boat that moves at 12 m/sec on still water is on a river whose current flows at 3 m/sec; the river is 3 km wide. (a) How long will it take the boat to cross the river and return? (b) How long will it take it to go 3 km upstream and return?

3. Consider a river 10 m wide that flows 30 km/sec, and a boat that travels 3×10^8 m/sec. (a) How long will it take the boat to cross the river and return? (b) How long will it take it to go 10 m downstream, and then back?

4. Show that, for the earth supposedly traveling 30 km/sec with respect to the hypothetical ether, the "cross-stream" and the "up-and-down-stream" light beams should arrive with a time difference equal to 5×10^{-9} of the total travel time of the light in the interferometer.

(18-3)

5. A future spaceship, receding from the earth at half the speed of light, fires from its nose a rocket which travels at half the speed of light with reference to the ship. With what speed does the rocket travel with reference to the earth?

6. A beam of ions is shot from the nose of a spaceship at a speed of $0.9c$. The ship is leaving the earth at $0.6c$. What would earthly observers measure the speed of the ions to be?

7. From the rear of the ship of Question 5, a ray of light is directed backward toward Earth. What will earthly observers find the speed of the light ray to be? (Here, $v = -c$.)

8. In Question 6, the ion beam is directed backward toward the earth. What is the speed of the beam, as observed by scientists on the earth?

9. Imagine a spaceship which has a speed of $0.6c$ relative to the earth. Its rest length is 100 m. What would we measure its length to be as it passed by?

10. A spherical asteroid 10 km in diameter passes the earth with a relative velocity of $0.4c$. What do astronomers measure its diameter to be in the direction of its motion? (This is a purely fictional asteroid—real ones have *much* smaller speeds.)

11. Intercontinental planes will no doubt be traveling at 3600 km/hr in the near future. By how much would such a plane appear to be shortened, as observed from the ground? Take its rest length to be 150 m.

12. Consider the asteroid of Question 10 to be moving at a more realistic relative velocity of 40 km/sec. By how much would astronomers observe its diameter to be shortened in its direction of motion?

13. An electron has a rest mass of 9.11×10^{-28} gm. What is our measure of the mass of an electron traveling at $0.99c$?

14. The rest mass of a proton is 1.67×10^{-24} gm. What is our measure of the mass of a proton whose velocity is $0.9c$?

15. By what fraction is the mass of the earth increased (from the viewpoint of an observer stationary relative to the solar system) due to its 30 km/sec orbital speed? What does this amount to, in metric tons? ($m_0 = 6 \times 10^{27}$ gm.)

16. By what fraction will the rest mass of a molecule be increased when it is in the 2500-m/sec blast of a rocket exhaust?

17. A free neutron (not in an atomic nucleus) has an average lifetime of about 1000 sec. How fast must a beam of neutrons be traveling for them to have a lifetime twice this long?

18. A K^+ meson has an average lifetime of about 10^{-8} sec. How fast must it be moving to increase this average lifetime by 50 percent?

(18-5)

19. In Mr. Tompkins' city, how long a time (by his own watch) would it take a cyclist to go 1/2 mi at 15 mi/hr? (Distance measured by counting blocks, which are 16 per mile, measured on the ground.)

20. To the cyclist of Question 19, how many feet long would a city block appear to be?

(18-6)

21. A space traveler is 30 years old when he leaves to explore the Galaxy on a ship traveling at 2.4×10^8 m/sec. He returns 50 years later (by Earth calendars). About how old does the traveler appear to be?

22. In the year 2010, a traveler leaves Earth on an exploratory voyage at a speed of $0.99c$. His own accurate clock-calendar tells him that his trip has lasted just 10 years when he arrives back on Earth. What do the Earth calendars say the year is?

(18-7)

23. Confirm the figures given on p. 333 for the flashlight radiation and the heated water.

24. If a block of iron (specific heat = 0.1 cal/gm/°C) has a mass of exactly 1000 kg at 0°C, by how much will its mass be increased when it is heated up to 100°C?

25. A nuclear reactor-powered generating plant produces electric power at an average rate of a million watts. The energy is produced by the conversion of the mass of the nuclear fuel. By how much will the mass of the fuel have been reduced at the end of a year?

26. The sun loses about 4×10^6 metric tons of mass per second, by radiating the equivalent energy into space. What is the wattage of the sun?

27. The earth's orbital speed is about 30 km/sec. (See Question 15.) What is the earth's orbital KE (from KE = $1/2\ mv^2$)? What mass does this amount of energy correspond to? Does this agree with the answer to Question 15?

chapter / nineteen

The General Theory of Relativity

19-1 Acceleration and Gravity

One of the basic postulates of Einstein's Special Theory of Relativity is that it is senseless to speak about absolute motion through space and that only the relative motion of one system with respect to another system can be considered as a physical reality. An observer enclosed inside a windowless vehicle moving with a constant velocity has no way of telling whether he is in a state of motion or at rest, no matter what kind of physical experiments (mechanical, optical, electric, or magnetic) he performs inside his enclosure in order to answer this question.

But what about accelerated motion? When an airplane speeds up along the runway prior to the takeoff, you are pressed to the back of your seat, and it is not necessary to look through the window to know that you are subjected to an acceleration. If the flight is smooth, the conditions inside the airplane are exactly the same as if it were standing at the airport, but any accelerations caused by air currents in bad flying weather are certainly very noticeable to the passengers. Does this mean

that, whereas one should not talk about absolute velocities, absolute accelerations have a definite physical meaning?

This problem was attacked by Einstein about 10 years after his original publication of the theory of relativity (now known as the *Special Theory of Relativity*), and its solution (published in 1916) resulted in further theoretical developments now known as the *General Theory of Relativity*. He showed that ***all the physical phenomena that are observed within an accelerated system are identical with those occurring in a resting system placed in a gravitational field.*** To understand this idea, let us consider events taking place in the passenger cabin of a rocket ship traveling through space, far away from any gravitating celestial bodies. (In his original paper, Einstein spoke about a box being accelerated through space by the pull of a rope attached to it.) If the rocket motors are shut off, the ship coasts freely through space with constant velocity (Newton's first law), and the conditions in the cabin are the same as those described by the famous French fiction writer Jules Verne in his novel *A Trip Around the Moon*. The travelers and all the objects within the cabin are floating freely in the air, since there is no gravitational force to pull them in any direction (Fig. 19-1A).

Suppose now that the motors are started so that our rocket ship begins to gain speed. Since the velocity of the ship is now increasing, the things floating in the cabin will continue to move with the same velocity with which the ship was originally coasting through space, and they all will be collected at the rear wall of the cabin and pressed to it by the force of the acceleration (Fig. 19-1B). Realizing the situation, the travelers will rise to their feet and will stand on the rear wall of the cabin as if it were the floor of their laboratory back on the earth (Fig. 19-1C). Being of scientific mind and knowing that the earth is far away, they may try to perform some experiments in order to find out what the difference is between the force acting on them because of the operation of the rocket motors and the force of gravity acting on the earth's surface. Lifting a wooden ball, which for some unknown reason is among the other things in the cabin, one of the travelers (let us call him Dr. A) feels that it presses against his hand as though attracted toward the "floor" of the cabin.

"Feels exactly as if I am holding a wooden ball back home," says Dr. A. "If I let it go, it will drop down as it would on the earth."

"Yes, *as if*!" counters Dr. B. "But, you know very well there is no gravity here and that the ball will simply approach the rear wall of our cabin because, once you let it go, it will continue to move with a constant velocity while our rocket is accelerating and gaining speed."

"What's the difference?" objects Dr. A. "It falls just the same. Is there any way to check whether its fall is due to the acceleration of our rocket ship or to the gravitational attraction of some very large mass under our feet?"

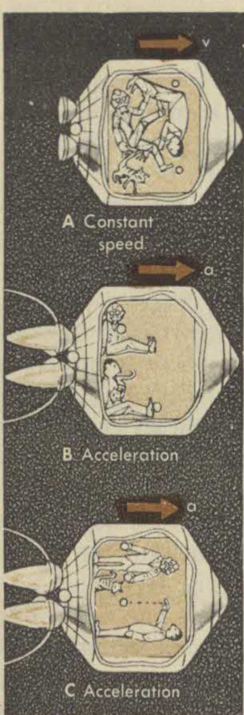


FIG. 19–1 Events taking place in an unaccelerated rocket ship (A); and in an accelerated one (B and C).

"Why don't we repeat Galileo's experiment, dropping a light and a heavy ball simultaneously and see if they fall down at the same speed," suggests Dr. B.

"Very well, here is an iron ball; God knows how it got here," says Dr. A, holding the two balls in his hands. "Now, I let them go!"

You can readily imagine what happens when Dr. A releases the two balls. They will, of course, move *side by side* with the same velocity that the rocket ship had at the moment of their release. The "floor" of the cabin, constantly gaining velocity, would finally overtake them, and hit them both at the same time. From the point of view of the rocket-ship travelers, however, it would seem that both balls move toward the floor and hit it simultaneously. This is, of course, the same result as was obtained by Galileo on the Tower of Pisa, but whereas the statement that "all material bodies fall down with the same speed" remained an unexplained puzzle for centuries, the side-by-side motion of the two balls in an accelerated rocket ship is a simple consequence of the law of inertia.

The principal point of Einstein's paper was that the behavior of material objects in an accelerated system and their behavior in a gravitational field are not only similar to one another but absolutely identical. In other words, ***no matter what kinds of physical experiments we carry out in a closed cabin, we can never find out whether the cabin is resting on the surface of some massive planet or is being accelerated through space far away from any gravitating masses.***

19-2 Gravitational Deflection of

One characteristic of a good physical theory is that it not only explains known facts but is also able to make new predictions that can be checked by experiment. Such was the case with Einstein's views concerning the identity of acceleration and gravity. Let us return to our rocket ship and see what would happen if the two scientists tried to perform some optical experiments inside their cabin. One of the basic properties of light is that it propagates along straight lines, so that, if we direct a beam of a flashlight at the wall, the illuminated spot will be directly opposite the source. If one repeats this experiment in an accelerated rocket, however, the situation will be rather different, as is shown in Fig. 19-2. If the rocket were not accelerated, the beam of light coming from the source on the left would propagate straight across the cabin to produce luminous spots S_1, S_2, and S_3 on the transparent plates of fluorescent glass (No. 1, No. 2, and No. 3) placed across its way. If, however, the rocket is accelerated, the luminous spots will no longer lie on a straight line. The light will take equal intervals of time to cover the

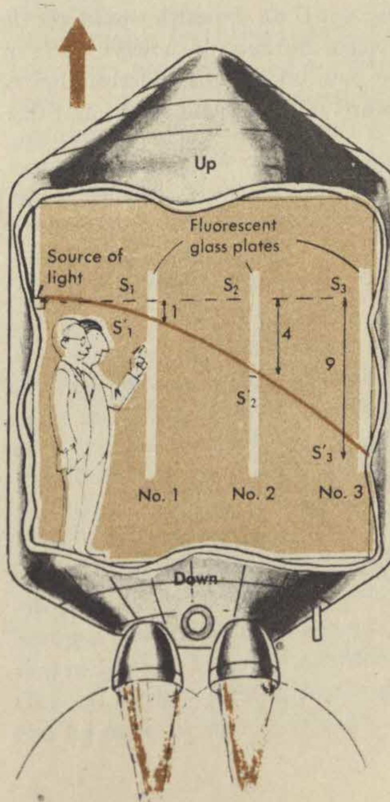


FIG. 19-2 The deflection of a beam of light as observed in an accelerated rocket ship.

distances between Source and No. 1, between No. 1 and No. 2, and between No. 2 and No. 3. Because of the accelerated motion of the rocket the displacements of the screens during these time intervals will stand in the ratios $1^2 = 1$, $2^2 = 4$, $3^2 = 9$. Thus the spots S_1', S_2', and S_3', tracing the motion of the light beam will lie on a *parabola* rather than on a straight line. Therefore, the observers in the cabin, considering themselves to be in a gravitational field, will be under the impression that the light beam is deflected by the gravitational field in the same way that a flying bullet is. If the analogy discussed above between the motion of material bodies in an accelerated system and in a gravitational field is more than an analogy and if the two things are really identical, we should expect that ***light rays propagating through any gravitational field should be deflected in the direction of the acting force.***

This conclusion was first tested in the observation of a total eclipse of the sun in Africa by a British expedition in 1919. These observations proved that light rays from distant stars that pass close to the massive body of the sun are indeed deflected. The situation is shown schematically in Fig. 19-3. If the sun were in some other part of the sky, so that the starlight propagated along straight lines (the dashed lines in Fig. 19-3), an observer O on the earth would get the value θ for the angular distance between the stars. If, however, the rays coming from the two stars were to pass on opposite sides of the sun, their paths would be deflected toward the sun (solid lines), and the observed angle θ' would

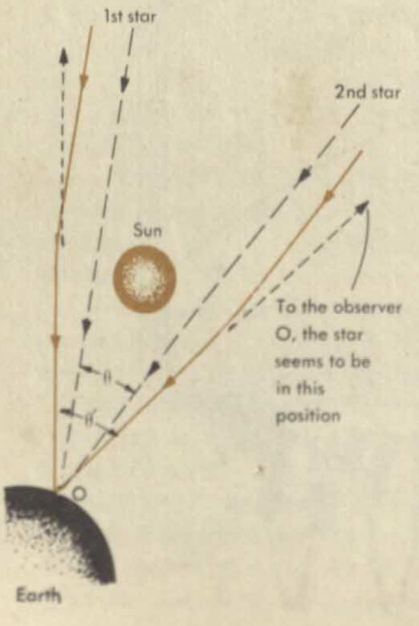


FIG. 19-3 The deviation of light from a star when the light passes close to the sun.

be larger. Thus the gravitational field of the sun acts as though it were a converging lens, making the distance between the two stars look larger than it actually is. The test could be made, of course, only during a total solar eclipse, since otherwise the light from the stars would be lost in the brilliant sunshine. The observation of the 1919 eclipse and other later observations confirmed the expected bending of light rays in the gravitational field of the sun.

The deflection of starlight as it passes near the sun is very small, however, and is difficult to measure with much precision. In recent years, several other theories (similar to Einstein's General Theory of Relativity, but differing from it in some details) have been proposed. So far, the experimental data from eclipse photographs have not been precise enough to show that any one of the theories is right and the others wrong.

19-3 Other Consequences of the General Theory of Relativity

Einstein's ideas concerning the nature of the gravitational field also predicted other consequences which could be confirmed by astronomical observations. One of them pertained to the motion of planets around the sun.

The major, i.e., longest, axis of the elliptical orbit of the planet Mercury does not remain always fixed in the same direction, but rather rotates, or *precesses*, about the sun approximately 575 seconds of arc per century (Fig. 19-4). Most of the precession can be explained by purely Newtonian mechanics, applied to the gravitational pulls of the other planets as they, too, revolve around the sun and tug at Mercury as they pass. When all these effects are calculated, 43 sec/century still remains unaccounted for.

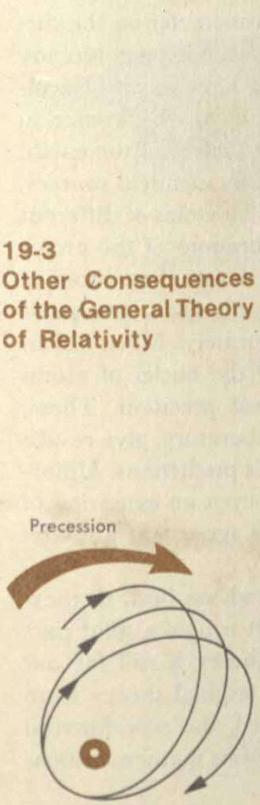


FIG. 19-4 Precession, or rotation, of an elliptical orbit.

Einstein's General Theory of Relativity predicts results a little different from Newton's, especially for a planet like Mercury, which has a very elliptical orbit and is in a very strong gravitational field close to the sun. Einstein's calculations showed that the major axis of Mercury's orbit should be expected to precess at the rate of just 43 sec/century, even without the pulls of any other planets. The General Theory of Relativity thus gave the answer to a puzzle over which astronomers had been scratching their heads for many decades.

But astrophysicists cannot let well enough alone; it has recently been noted that if the sun were not quite perfectly spherical, it would have an effect on Mercury's precession that had not previously been taken into account. Experiments are now going on to determine whether the rotation of the sun bulges its equator enough to put the predictions of the General Theory of Relativity in question. Eventually, if measurements can be made accurately enough and if we learn more about how the

mass of the sun is distributed in its deep, rotating interior, it may be possible to choose between Einstein's General Theory of Relativity and some of the competing not-quite-the-same theories.

Another important consequence of the General Theory of Relativity was that gravity should influence the rate of all physical processes by slowing them down. Thus, on the surface of the moon, where gravity is weaker than on the surface of the earth, a chronometer should gain time with respect to an identical terrestrial chronometer, whereas on the surface of the sun, where gravity is much stronger, it should trail behind. Of course, it is impossible to place a man-made chronometer on the surface of the sun, but fortunately there are natural chronometers already there: the atoms which tick their time by emitting light waves of well-defined frequencies. Thus, in order to see if there is any difference in the clock rate on the surface of the sun and on the surface of the earth, one should compare the frequency of light emitted by identical sources, one on the sun and one on the earth. Light emitted by atoms of different elements provide means for such a study. The vibrations of the atoms in the sun's strong gravitational field should be slower than those of similar atoms on the earth. But this difference is only about two parts in a million and is very difficult to measure accurately. More recent experiments have used the vibrations emitted by the nuclei of atoms (gamma rays), which can be measured with great precision. These, measured at different elevations within the same laboratory, give results that (like the shift in sunlight) agree with Einstein's predictions. Unfortunately, however, all of the competing theories—even an extension of Newton's laws—predict the same thing, so that this agreement provides no basis for choosing among them.

Einstein's Special Theory has no competitors, and has been so thoroughly confirmed by experimental evidence that it is now a solid part of the foundations of physics. But the General Theory is still far out along the frontiers of science. Although it is the original theory from which all its present competitors have been derived, the experimental evidence is not yet good enough to say which one of its many modifications, if any, is correct.

19-4 Gravity and Space Curvature

The second great step made by Einstein in his General Theory of Relativity was to associate a gravitational field with a curvature of the four-dimensional space in the neighborhood of gravitating masses. What does it mean that space is curved? The best way to understand the cumbersome notion of curved space is through an analogy with curved surfaces that have only two dimensions and thus can be easily visualized. We know very well the difference between a plane surface, such as the surface of a table, and curved surfaces, such as those of a football or a saddle. Mathematicians distinguish between two kinds of curved surfaces: those with *positive curvature* and those with *negative curvature*.

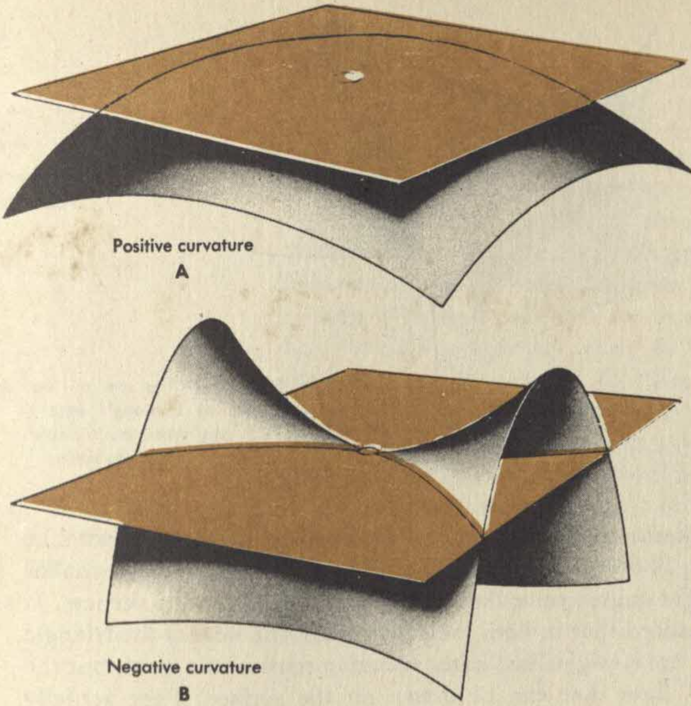
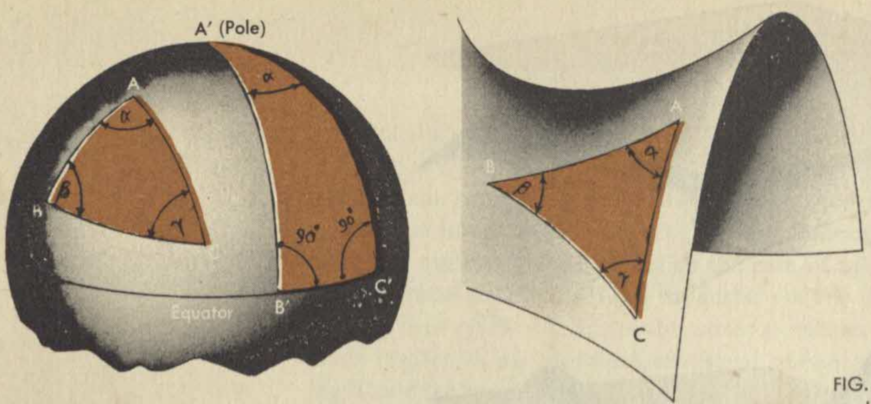


FIG. 19-5 Surfaces of (A) positive; and of (B) negative curvature. Note that in the case of positive curvature the entire curved surface is on the same side of the tangent plane, while in the case of negative curvature, part of the curved surface is on one side of the tangent plane and part is on the other side.

In order to tell which is which, we draw a plane tangent to the curved surface at one point. If the curved surface lies entirely on one side of the plane (Fig. 19-5A), the curvature of the surface is called positive. If, on the other hand, the plane and the surface intersect so that one part of the curved surface lies above and another part below the plane (Fig. 19-5B), the curvature is called negative. We can easily see that, according to this definition, the surface of a sphere or an ellipsoid has a positive curvature, while the curvature of any saddle-shaped surface is negative.

The difference between a surface of positive and one of negative curvature shows in the properties of geometrical figures drawn on them. While figures drawn on a plane surface are subject to the rules of classical Euclidean geometry, it is not so for figures drawn on curved surfaces. Consider, for example, the theorem of Euclidean plane geometry according to which the sum of the angles of a triangle is always equal to two right angles. This theorem does not hold for spherical triangles formed by arcs of great circles connecting three points A, B, and C on the surface of a sphere (Fig. 19-6A), since the surface, so to speak, "bulges up" between the vertices of the triangle. The simplest way to see this is to consider a triangle $A'B'C'$ with one vertex at the pole and two others on the equator of the sphere. Since the meridians forming two sides of that triangle make right angles with the equator, the sum of the angles $A'B'C'$ and $A'C'B'$ is already equal to two right angles, and we have in addition the third angle $B'A'C'$ which can be quite large. An opposite



$\alpha+\beta+\gamma > 180°$

A

$\alpha+\beta+\gamma < 180°$

B

FIG. 19-6 The sum of the angles of a triangle equals 180° only when the triangle is drawn on a plane surface.

result will be obtained in the case of a triangle drawn on a saddlelike surface (Fig. 19-6B). The sum of the angles of a triangle there is smaller than two right angles, since the surface "sinks" between the vertices. It should be noticed that in both the above cases, the sides of the triangle are actually not straight lines in the common sense of the word, but the "straightest" lines that can be drawn on the surface. They actually represent the *shortest distances* between the two given points, and are known in mathematics as *geodesic* lines, or simply *geodesics*. In the geometry of curved surfaces, geodesics play the same role that straight lines do in ordinary plane geometry.

From the simple facts concerning the geometry of surfaces with only two dimensions, we can now generalize for the case of three-dimensional space. Since we are ourselves three-dimensional beings, and hence cannot look at the space we live in *from the outside* as we can look at the surfaces, we are deprived of the ability to visualize curved spaces as easily as we can curved surfaces. What we can do, however, is to say that the three-dimensional space which surrounds us on all sides is curved if its geometrical properties deviate from the laws formulated by Euclid. Thus, if the sum of the angles of a triangle formed by any three points in space is larger than two right angles, we ascribe to the space a positive curvature; if the sum of the angles is less than 180°, the curvature of space is considered to be negative.

In order to give physical meaning to these rather abstract considerations, imagine three astronomers located at three observation points (planets, or even spaceships) around the sun (Fig. 19-7). Each of the three astronomers uses a very accurate instrument for measuring the angles of the triangle *ABC* formed by the three positions in space. If the sun were not inside this triangle, light rays would propagate along "conventional" straight lines (broken lines in Fig. 19-7), and our astronomers would confirm the classical Euclidean statement. The presence of the gravitational field of the sun, however, will cause light rays to be

bent (solid lines in Fig. 19-7), and adding together their measurements of the three angles, our astronomers will get a value larger than two right angles. Einstein's revolutionary idea was to ascribe the finding of the three astronomers in the above hypothetical case, *not to the deflection of light rays propagating through Euclidean space, but rather to the deviation of the geometry of the space itself from the classical Euclidean rules*. In other words, *light rays always propagate along the "straightest" (geodesic) lines, and the deviation from the rules of Euclidean geometry obtained by measurements carried out in the neighborhood of the sun is due to the curvature of space caused by the presence of the sun's large gravitating mass*. Since the sum of the angles of a triangle is in this case larger than two right angles, the curvature of space in the neighborhood of the sun is to be considered as positive. The idea of replacing the picture of a physical deflection of light rays propagating through a gravitational field by a change in the geometry of space caused by the presence of the gravitating mass turned out to be very helpful to the understanding and mathematical description of the phenomenon of gravity.

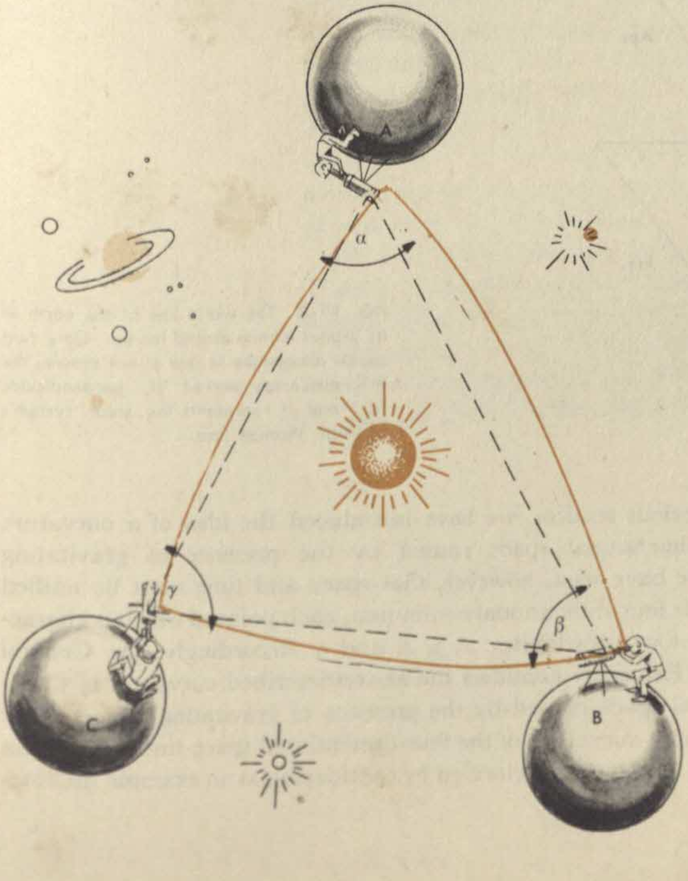


FIG. 19-7 Three observers checking Euclidean geometry for a triangle described around the sun.

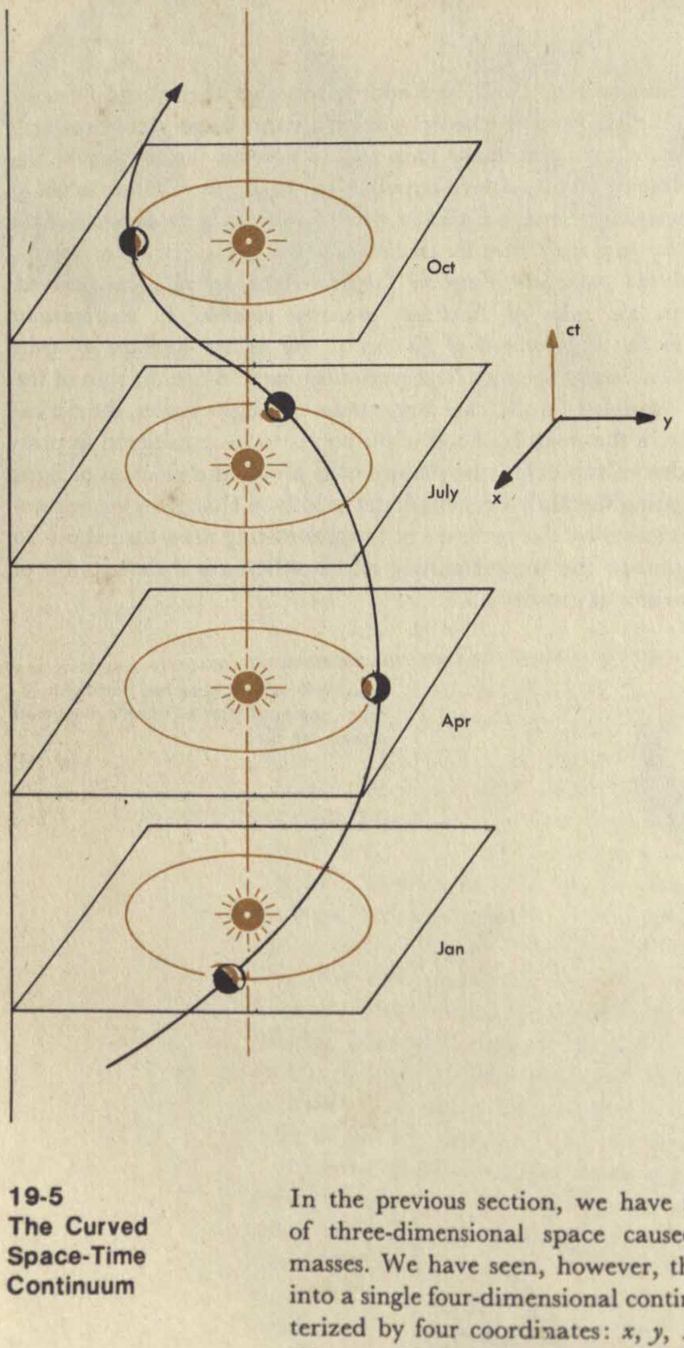


FIG. 19–8 The world line of the earth in its orbital motion around the sun. Only two space dimensions (x and y) are shown; the third dimension (marked "ct," perpendicular to x and y) represents the solar system's passage through time.

19-5 The Curved Space-Time Continuum

In the previous section, we have introduced the idea of a curvature of three-dimensional space caused by the presence of gravitating masses. We have seen, however, that space and time must be unified into a single four-dimensional continuum, each point of which is characterized by four coordinates: x, y, z, and t. Accordingly, the General Theory of Relativity considers the above-described curvature of three-dimensional space caused by the presence of gravitating masses to be the result of a curvature of the four-dimensional space-time continuum itself. This notion can be clarified by considering as an example the rota-

tion of the earth around the sun. Since the earth's orbit lies in a plane, we need retain only the two space coordinates x and y located in this plane and replace the third space coordinate z by time t. It is convenient to multiply time by the velocity of light c, so that this coordinate will have the same physical dimensions as the simple space coordinates x and y. As a result, we obtain the picture shown in Fig. 19-8, in which each plane perpendicular to the ct axis gives the position of the earth in its orbit at the time corresponding to the position of the plane. Connecting the consecutive positions of the earth by a continuous line, we obtain a helix that winds around the time axis passing through the sun. Such lines, which if we include the third z coordinate of space would run through the four-dimensional space-time continuum, are known as the *world lines* of the material bodies whose motion they describe.

In our example, the world line of the sun is a straight line perpendicular to the xy plane, whereas the world line of the earth is a helix winding around it. Such would be the situation if we considered the space-time continuum to be subject to the rules of Euclidean geometry. In the General Theory of Relativity, however, the four-dimensional space-time continuum is itself considered to be curved, so that the conventional straight lines of Euclidean geometry must be replaced by geodesic lines in curved four-dimensional space. It was shown by Einstein that the helical world line of the earth shown in Fig. 19-8 is actually such a geodesic, or "straight line" in the curved space surrounding the sun. Thus, just as in the case of the deflection of light rays passing close to the sun, *the orbital motion of the earth can be interpreted, not as being caused by a certain physical force exerted on it by the sun, but as the result of the curvature of space in the sun's neighborhood.*

We can summarize this section by saying that Einstein's General Theory of Relativity gives a geometrical interpretation to the Newtonian theory of universal gravity: *instead of saying that the propagation of light and the motion of material bodies are compelled to deviate from straight lines by the force of gravity, we say that this motion takes place along the "straightest" (geodesic) lines in a space that is curved by the presence of gravitating masses.*

chapter / twenty

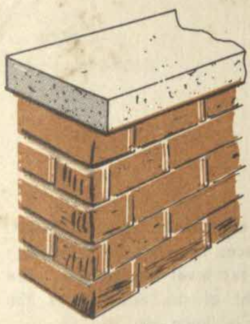
The Molecular Nature of Matter

20-1 The Molecular Hypothesis

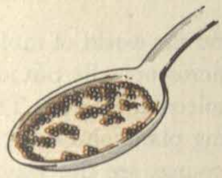
Surveying the physical properties of different substances encountered in nature, we find a great deal of variety. Some of the substances are normally solid, melting and turning into gas only at extremely high temperatures. Others are normally gaseous, becoming liquid and freezing only when the temperature drops close to absolute zero. Some liquids are of high fluidity while others are very viscous. Some substances, generally known as metals, possess a high degree of electric and thermal conductivity, while others, the dielectrics, are very good insulators. Some substances are transparent to visible light while others are completely opaque; some possess a high refractive index and some a low one. . . .

We ascribe all these differences between substances to differences in their internal structure and attempt to explain them quantitatively as well as qualitatively as being due to different properties and interactions of the structural elements of matter. We assume that such seemingly homogeneous substances as air, water, or a piece of metal are actually

composed of a multitude of extremely small particles known as *molecules*. All molecules of a given pure substance are identical, and the differences in physical properties between various substances are due to the differences between their molecules. There are as many different kinds of molecules as there are different substances, of which there is indeed an enormous number. There are the molecules of oxygen and of mustard gas, the molecules of water, alcohol, and glycerine, the molecules of iron, asbestos, and camphor, the molecules of gelatine, insulin, and fats. . . .



A a solid



B a liquid

C a gas

FIG. 20-1 Three states of matter.

The molecules forming any given material body are held together by *intermolecular forces*, which are determined by the nature of the molecule. These forces resist the tendency of internal thermal agitation to break up the molecular aggregates. If intermolecular forces are strong, molecules will be as rigidly cemented together as the bricks in a garden wall (Fig. 20-1A), and the material will remain solid up to very high temperatures. If these forces are comparable to the forces of thermal agitation, they may not be able to hold the molecules rigidly in their places and may permit the molecules to slide more or less freely past each other as if they were grains of fresh Russian caviar (Fig. 20-1B). Thus the substance will keep its volume, but it will take the shape of the container in which it is placed. The viscosity of the liquid will depend on how easily this sliding of molecules can take place, and the substance will become more and more fluid as its temperature and thermal agitation rise. If the intermolecular forces are very weak, the molecules will fly apart in all directions and the material must be kept in a closed container like a bunch of agile flies in a glass jar (Fig. 20-1C). This picture explains the high compressibility of gases, since compression results only in the reduction of the free space in which the molecules are moving. We can get an idea about the amount of free space between the molecules of a gas by comparing the density of the gas with its density in the liquefied state, when the molecules are packed together. For example, the density of atmospheric air, under normal conditions, is 0.0012 gm/cm^3, and the density of liquid air is 0.92 gm/cm^3 or 800 times larger. Since, in the liquid state, the distances between the centers of neighboring molecules are approximately equal to their diameters, the distances in the gaseous state must be $\sqrt[3]{800} = 9.3$ times the diameters. This relation is shown in Fig. 20-2.

20-2 Brownian Motion

Although molecules are too small to be seen individually even through the best microscope, their thermal agitation can be noticed by observing the movement of small particles of smoke floating in the air. This phenomenon is called *Brownian motion*, after the British botanist Robert Brown (1773–1858), who in 1827 observed the irregular motions of plant spores floating in water. Such small particles play the role of an

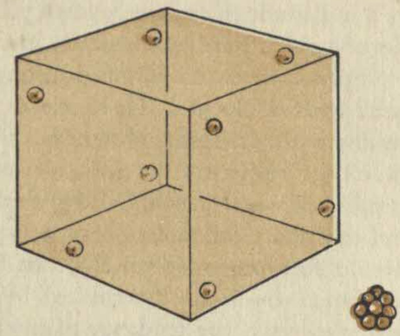


FIG. 20–2 The approximate relative diameters and distances between air molecules in the normal sea-level atmosphere. The blob to the right of the box shows the same molecules in liquid air.

intermediary between our familiar surroundings and the world of molecules, since they are large enough to be observed microscopically but are also sufficiently small to be affected by irregular molecular motion. The situation is similar to that of a pilot in a high-flying plane observing a Navy task force in a choppy sea. He cannot, of course, see the waves themselves, and tiny life rafts that follow every movement of the water are invisible to him; at the other extreme, the big aircraft carriers will seem to float without any disturbance at all. But the medium-sized ships, which he can still see, will show a definite roll, and our pilot will know that the sea is rough. In Fig. 20-3, we show the successive positions of a smoke particle, 1 micron ($= 10^{-6}$ m) in diameter, dancing its Brownian dance in atmospheric air. The study of Brownian motion by both Einstein and the French physicist Jean Perrin (1870–1942) led to indisputable proof of the reality of the thermal motion of molecules and gave us valuable information concerning the amount of kinetic energy involved in it.

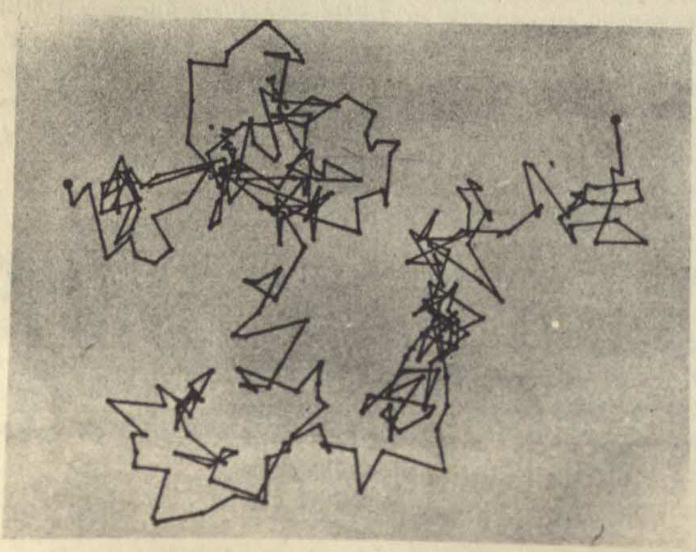


FIG. 20–3 Successive positions of a smoke particle in air, recorded at one-minute intervals (according to J. B. Perrin).

Figure 20-3 is not an actual detailed chart of the real motion of a smoke particle. As the caption implies, the position of the particle (on a magnified scale) has been recorded at the end of each minute, and then these points have been joined by straight lines. The actual path of the particle between any two points was itself a jagged, irregular one similar to the larger-scale path of Fig. 20-3, but on a scale too small to measure and record. The point of the matter is that if statistical methods are applied to data such as those recorded in the illustration, it is possible to calculate the *average velocity* of the particle during its erratic journey. If the mass of the Brownian motion particle is known, it is then a simple matter to get its *average* kinetic energy, from $\frac{1}{2}mv^2$.

But what is the use of this bit of knowledge, other than a little intellectual exercise? The smoke particle is enormously larger than the molecules of the air in which it is suspended. Its motion, however, is caused by millions of collisions per second with the air molecules. One individual collision will not move the particle appreciably; but occasionally the number of air molecules striking it on the left side, say, will happen to be greater than the number striking it on the right side, and as a result the particle will move a visible distance to the right. Here again, a complicated statistical analysis of all possible collisions leads to the theorem of the *equipartition of energy*. This theorem states that ***in the random motion of a large number of colliding particles, the average kinetic energies of all particles are the same, regardless of their masses.*** In other words, on the average, $\frac{1}{2}m_P v_P^2$ for the large particle is exactly equal to $\frac{1}{2}m_M v_M^2$ for a molecule of air.

Brownian particles in the neighborhood of 1 micron in diameter are observed to move at room temperature with an average velocity of 0.65 cm/sec, and since their mass is about 5×10^{-13} gm, their kinetic energy must be about 10^{-13} erg. Thus, by the equipartition theorem, this value must also represent the kinetic energy of individual air molecules at this temperature.

With an increase of temperature, the intensity of Brownian motion also increases, so that by direct observation of the tiny particles suspended in a gas, or a liquid, we can study how the energy of thermal motion depends on temperature. The results of such experiments performed between 0° and 100°C are shown in Fig. 20-4. The observed points are located on a straight line and, extrapolating it in the direction of lower temperatures, we come to the conclusion that Brownian motion must completely stop at −273°C.

We recall that −273°C is 0°K, the zero of the absolute temperature scale (see Fig. 10-4). From this, and from the straight-line relationship in Fig. 20-4, we must come to the conclusion that ***the average kinetic energy of the thermal motion of molecules is directly proportional to the absolute temperature.***

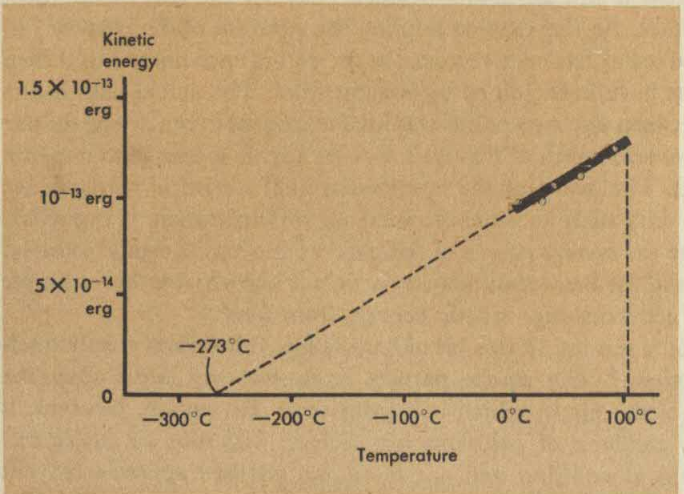


FIG. 20–4 A schematic diagram showing the kinetic energy of Brownian motion measured for temperatures between 0° and 100°C, and extrapolated to lower temperatures.

As we deduced above from the Brownian motion of smoke particles, the average kinetic energy of an air molecule at room temperature (about 300°K) is about 10^{-13} erg. Now we may imagine taking an air molecule at absolute zero and warming it up to room temperature a degree at a time. Since its 300°K energy is 10^{-13} erg, this means we would have to add $10^{-13}/300 = 3.3 \times 10^{-16}$ erg to raise the temperature of 1 molecule of air by 1°C.

On the other hand, we can deal with air in wholesale quantities, paying no attention to the billions of molecules which it contains. From such bulk experiments, we know that the specific heat of air is 0.16 cal/gm/deg. Converting this to ergs, we have

$$0.16 \text{ cal} \times 4.18 \text{ joules/cal} \times 10^7 \text{ ergs/joule} = 6.7 \times 10^6 \text{ ergs}.$$

It thus takes 6.7×10^6 ergs to raise the temperature of 1 gm of air by 1°. Then, merely by the simple division of $6.7 \times 10^6/3.3 \times 10^{-16}$, we learn that 1 gm of air must contain 2×10^{22} molecules; or in other words that an air molecule has a mass of $1/(2 \times 10^{22})$ = about 5×10^{-23} gm. (This, of course, is a rough average for the 80 percent nitrogen and 20 percent oxygen that are the main constituents of air.)

Since 1 gm of liquid air occupies a volume of about 1 cm^3 (its density is 0.92, i.e., about 1), the volume of a single air molecule must be about $1 \text{ cm}^3/(2 \times 10^{22}) = 5 \times 10^{-23} \text{ cm}^3$, and its diameter about $\sqrt[3]{5 \times 10^{-23}} = 4 \times 10^{-8}$ cm.

We can also use the data above to estimate the velocity of molecular motion. At room temperature, we have found the average energy of a molecule to be 10^{-13} erg, and its mass to be not far from 5×10^{-23} gm. Then, from $KE = \frac{1}{2}mv^2$,

$$10^{-13} = 0.5 \times 5 \times 10^{-23} v^2$$

$$v^2 = 10^{-13}/2.5 \times 10^{-23} = 4 \times 10^9$$

$$v = \sqrt{4 \times 10^9} = \sqrt{40 \times 10^8} = \text{about } 6 \times 10^4 \text{ cm/sec}$$

$$= \text{about } 0.6 \text{ km/sec.}$$

It is interesting to think that all this knowledge of invisible molecules has been derived from measurements made entirely on bulk material such as smoke particles and grams of air. We can summarize the results in Table 20-1:

TABLE 20-1 APPROXIMATE PROPERTIES OF AIR MOLECULES

Mass	5×10^{-23} gm
Diameter	4×10^{-8} cm
Velocity (at room temperature)	0.6 km/sec

20-3 Molecular Beams

Our faith in the existence of molecules, however, might be a little more firm if it rested on evidence somewhat more direct than used above. If we make a small hole in the wall of a vessel containing a considerable amount of gas, and the gas escapes out into the surrounding vacuum, the picture will look in general like that of a panicked crowd rushing out of a burning theater. Although the resulting stream of gas will be generally directed away from the opening, individual molecules inside the stream will continue to execute their irregular thermal motion, rushing in all directions and constantly colliding with each other (Fig. 20-5A). The motion of the gas streaming into the vacuum can be, in this case, considered as the motion of a continuous material, and it will be subjected to the laws of ordinary aerodynamics governing gaseous jets. The phenomenon will be different if the density of gas in the container is so small that the molecules have a very small chance of colliding with each other and hence can change their direction of motion only as the result of reflection from the walls of the container. In this case, the molecules will fly out of the container one by one (Fig. 20-5B); the rules of ordinary aerodynamics will no longer be applicable, and the velocity of the outgoing particles will be determined by the thermal velocity of their molecular motion, corresponding to the temperature of the gas. These *molecular beams*, in which the individual molecules go their own way and do not interact with the others, are very useful for the study of many molecular properties.

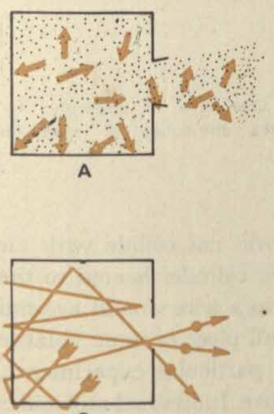


FIG. 20-5 Molecular beams: (A) when the density is high there are many collisions between molecules and the gas flows out in a continuous stream; (B) at very low density there are few collisions and the molecules escape as individuals.

The German physicist Otto Stern pioneered in the study of molecular beams, and, with his students and followers, developed several ingenious methods for directly measuring molecular velocities. One such piece of equipment is sketched in Fig. 20-6. The entire apparatus is in a high

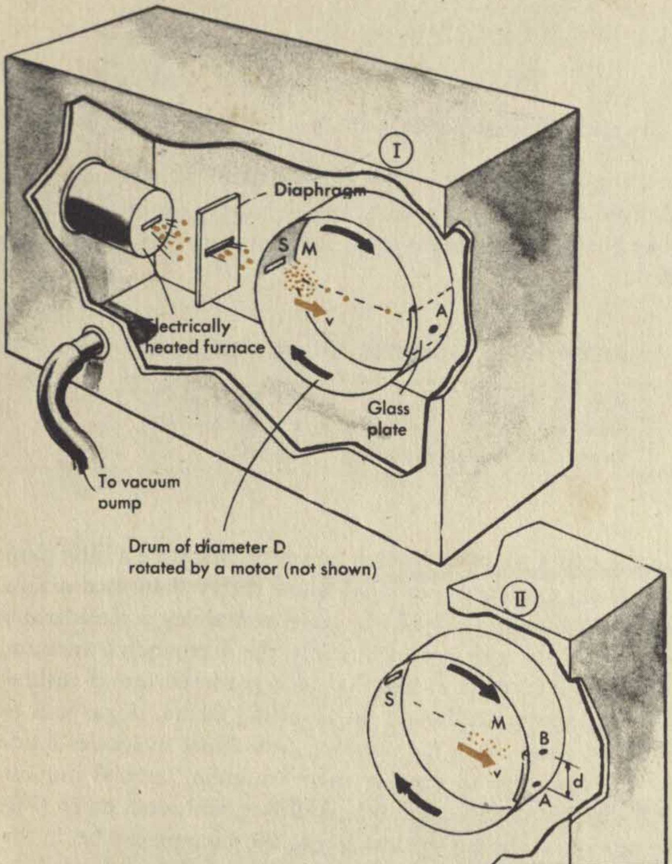


FIG. 20–6 One type of apparatus for measuring the distribution of molecular velocities.

vacuum so that the molecules being studied will not collide with air molecules. The source of the beam is a ceramic cylinder heated to the required high temperature by electric current in a wire wound around its surface. Within the cylinder is placed a small piece of some volatile material (sodium or potassium metals in this particular experiment), which gives rise to a gas of low density and pressure. Individual molecules of the gas fly out through a small opening at the base of the cylinder, and a thin molecular beam is cut out of the divergent stream by a slotted diaphragm.

This thin, parallel-sided beam of molecules falls generally on the outside surface of a rotating metal drum. Once each revolution, however, a thin slit S in the drum passes across the beam, and a small spurt of molecules M passes through to the inside of the drum, as shown in I. Directly opposite the slit in the drum is a glass plate; if the drum were

stationary, the molecules would pass across and strike the plate at point A. The drum, however, is rotating rapidly, and during the time the spurt of molecules crosses the diameter of the drum D, the drum has moved so that the flying molecules strike the plate at point B instead of at A (diagram II).

Let us say that the drum is rotating at an angular speed of f rev/sec, which is $2\pi f$ rad/sec. We recall that the linear speed of any part of a rotating body is $r\omega$, so the linear speed of the glass plate is $(D/2) \times 2\pi f = \pi Df$. In a time interval t, the glass plate will have moved through a distance $d = D\pi ft$, where t will be the time required for the molecules to cross the diameter of the drum. If v is the speed of the molecules, this time will be equal to D/v. Substituting this value for the time into the equation for d, we get

$$d = \pi Df \times \frac{D}{v}$$

$$= \frac{\pi f D^2}{v}$$

or

$$v = \frac{\pi f D^2}{d}.$$

From the above, the experimenter knows that any molecules deposited at a distance d from point A must have traveled with a speed given by the last equation. The values of thermal velocities obtained by this direct method are in perfect agreement with the values obtained by the less direct considerations described in the previous section.

It was found that not all the molecules in the beam struck the plate at the same spot—in other words, that not all the molecules had the same velocity. It is quite easy to measure the relative thickness of the deposit at various distances along the plate and thus to figure the relative number of molecules moving at each different velocity. Although the main bulk of the molecules move at about the velocity which would be expected from the temperature, there are always some molecules that move considerably slower or considerably faster than the average. Curves showing the distribution of molecular velocities at two different temperatures are given in Fig. 20-7. These deviations from the average result from the statistical nature of molecular motion and the irregularity of molecular collisions which may occasionally almost stop a molecule in its tracks or else send it rushing on with an abnormally high speed. The statistical study of the velocity distribution of the molecules of a gas was carried out in the last century on a purely theoretical basis by the British physicist James Clerk Maxwell (mentioned before in connection with his work on electric and magnetic fields), who derived a mathe-

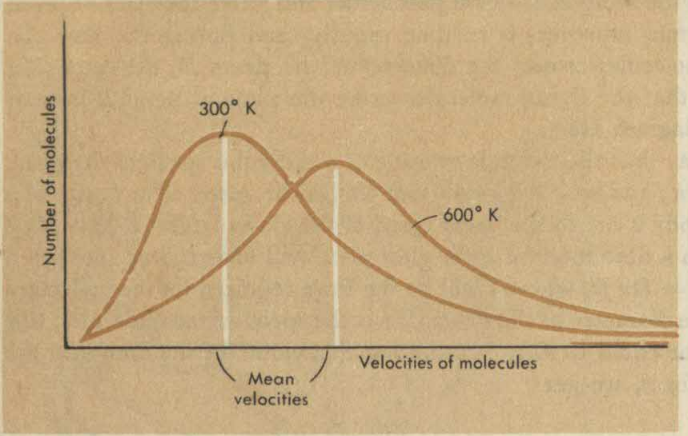


FIG. 20–7 Maxwell's calculation of the molecular velocities in a gas at two different temperatures.

matical expression describing the distribution of molecules of different velocities. The curve obtained by Stern in his experiments stands in excellent agreement with *Maxwell's distribution* of molecular velocities.

20-4 Kinetic Theory of Gases

The discussion in the previous section serves as the foundation for a mathematical theory known as *the kinetic theory of gases*. In Chapter 10, we dealt with the general gas law, $PV/T =$ constant, or $PV = cT$. This law can be interpreted as saying that for a given mass of gas, if the volume is kept constant, the pressure will vary directly as the absolute temperature, and that, for a constant temperature, the pressure varies inversely as the volume.

Let us see what comes of the assumptions on which an elementary form of the kinetic theory of gases is based. To avoid mathematical complications, we take all the molecules in a sample of gas as being identical, hard, perfectly elastic particles whose size is negligible, and which neither attract nor repel each other. (As a matter of fact, *all* these assumptions are wrong, but the ideal gas which they describe leads to a very simple mathematical treatment that gives a clear insight into the behavior of real gases. The deviations from actuality are quite small.)

Consider a closed vessel (Fig. 20-8) containing a certain amount of gas. The molecules of the gas are rushing ceaselessly in all directions, bounding off and colliding between themselves. The pressure of the gas on the surrounding walls is the result of the continuous bombardment to which the walls are subjected by the onrushing molecules. Suppose first that we keep the temperature of the gas (i.e., the velocity of its molecules) constant but reduce the volume. It is easy to see that in this case the number of molecules hitting a unit area A of the wall will

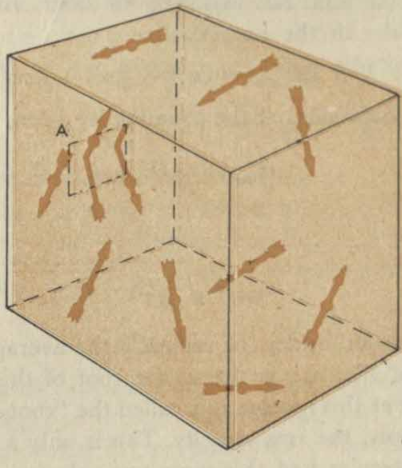


FIG. 20-8 Molecules of a gas bouncing (reflecting) from an enclosing wall, and thereby exerting a pressure on it. Consider the enclosure to be a cubical box measuring a cm on each side.

increase in inverse proportion to the decreasing volume. Since the pressure experienced by the walls is proportional to the total number of impacts they receive per unit time, we have here the explanation of the inverse proportionality between the pressure and the volume of gas.

Now keep the volume constant but increase the temperature and thus speed up the molecular motion. This will have two effects:

1. Since the molecules are moving faster, a larger number of them will collide with the wall area per unit time.
2. The impact of each molecule will be more violent because of its greater speed.

Since the pressure depends on both the number of impacts per second and the force of each impact, and since each of these factors is proportional to molecular velocity, the combined result is that the pressure will be proportional to the square of that velocity, i.e., to the kinetic energy of molecular motion, or, what is the same, to the absolute temperature of the gas.

The qualitative arguments given above will be more convincing if our conclusions are backed up with quantitative mathematical arguments as well. In the box of Fig. 20-8, molecules will be moving about at many different speeds, and the total energy of the molecules will be the sum of their individual kinetic energies:

$$\text{KE} = \tfrac{1}{2}mv_1^2 + \tfrac{1}{2}mv_2^2 + \tfrac{1}{2}mv_3^2 + \cdots$$

(We can use the same m for all the masses, since we have assumed the molecules to be identical.) There is some particular velocity, which we can call v_{rms}, such that, if all the molecules were going at this same rate

of speed, the total KE would be the same. In other words, if there are N molecules in the box, then

$$N \times \tfrac{1}{2}mv_{\text{rms}}^2 = \tfrac{1}{2}mv_1^2 + \tfrac{1}{2}mv_2^2 + \tfrac{1}{2}mv_3^2 + \cdots \tfrac{1}{2}mv_N^2.$$

Dividing both sides of the equation by $\frac{1}{2}Nm$, we get

$$v_{\text{rms}}^2 = \frac{1}{N}(v_1^2 + v_2^2 + v_3^2 + \cdots + v_N^2)$$

and

$$v_{\text{rms}} = \sqrt{\frac{1}{N}(v_1^2 + v_2^2 + v_3^2 + \cdots + v_N^2)}.$$

The expression under the radical is the average, or *mean*, of all N of the individual v^2's; v_{rms} is the square root of this mean of the individual squares. For this reason, it is called the "root-mean-square" velocity, or more simply, the rms velocity. This is only a special sort of average. If all N molecules had this same rms velocity, their total kinetic energy would be the same as the total kinetic energy of the N molecules actually moving with many different velocities.

We can now simplify matters by having all the molecules travel at this rms velocity, which is the molecular velocity associated with the given temperature. A further simplification can be made if we consider that at any instant, a third of the molecules, $N/3$, are bouncing back and forth vertically between top and bottom, another $N/3$ between the left face and the right face, and $N/3$ between the front and back sides; and all this without colliding with one another. (Actually, since we have assumed the molecules to be perfectly elastic, it makes no difference in the final answer if they *do* collide. However, if we avoid collisions it helps keep our mental picture a little clearer and simpler.) This seems very unreal and artificial, but a more sophisticated and realistic mathematical analysis brings us to exactly the same final result.

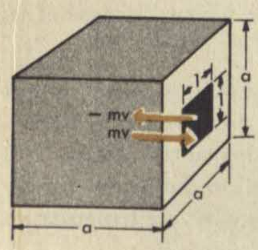


FIG. 20-9 The change in momentum of a gas molecule when reflected from a wall of its container.

Figure 20-9 shows one of the molecules bouncing off the right wall. Its original momentum was mv (the v, written without subscripts, means the rms velocity); after impact, it bounces back with momentum $-mv$, making a change in momentum of $2mv$. In order to find the pressure which results from many of these impacts, we must find how many molecules will strike per second on a unit area. A molecule going v cm/sec will require a/v sec to travel the a cm from one side of the box to the other and will therefore make v/a collisions per second with the two walls between which it is bouncing. On one wall, the collisions per molecule will thus be $v/2a$ per second. There are $N/3$ molecules bouncing between each pair of walls, so the total number of collisions on one wall is $Nv/6a$ per second.

It has been shown above that each collision produces a momentum

change of $2mv$, which means that the total momentum change per second on one wall is $2mv \times Nv/6a = Nmv^2/3a$. The wall has an area of a^2, so the momentum change per unit area per second is

$$\frac{1}{a^2} \times \frac{Nmv^2}{3a} = \frac{Nmv^2}{3a^3}.$$

From our earlier discussions of Newton's second law, we recall that force is equal to the rate of change of momentum, which is just what we have finished figuring. Hence the force on a unit area of wall, which is the pressure, is the last expression we have derived above:

$$P = \frac{Nmv^2}{3a^3}.$$

This expression tells us that the pressure is proportional to v^2 and thus is proportional to the average kinetic energy of the molecules, and also, by definition, proportional to the absolute temperature.

It is easy to reduce our expression for P to a somewhat more useful form. We can note two things in the expression: a^3 is the volume V of gas under consideration, which confirms Boyle's law that for any constant temperature, P is inversely proportional to V; and Nm is the total mass of the gas. Since the density d is Nm/V, the expression for pressure becomes simply

$$P = \frac{dv^2}{3}$$

or

$$v_{\text{rms}} = \sqrt{\frac{3P}{d}}.$$

A handbook tells us that the density of air at 1 standard atmosphere (1.013×10^6 dynes/cm^2) and 0°C (273°K) is 1.293×10^{-3} gm/cm^3. Substitution of these values into the equation gives

$$v = \sqrt{\frac{3 \times 1.013 \times 10^6}{1.293 \times 10^{-3}}} = 4.85 \times 10^4 \text{ cm/sec.}$$

The characteristic features of the thermal motion of gas molecules can be easily demonstrated by means of a gadget originally designed by Thomas B. Brown. The apparatus looks like a flat aquarium formed by two glass plates placed just far enough apart to allow Ping Pong balls, which represent the molecules, to be placed between them. At the bottom of the "aquarium," is a wooden cogwheel that can be driven at different speeds by an electric motor. The motion of this wheel agitates the Ping Pong balls and makes them move in an irregular fashion through the space between the bottom and the top of the container. When the wheel rotates very slowly, most of the balls remain at the bottom of the

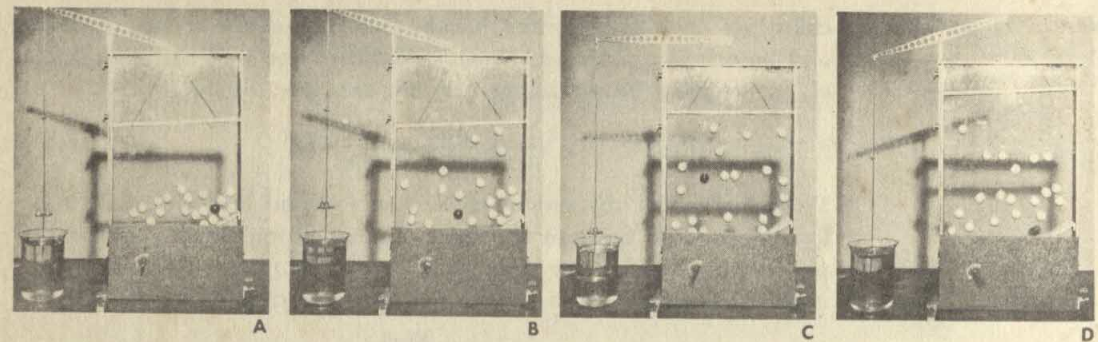


FIG. 20–10 A device for demonstrating the principle of the kinetic theory of gases. In (A) the drive-wheel rotates slowly (low temperature) and only a few of the "gas molecules" evaporate from the liquid state. (B), (C), and (D) represent increasingly higher temperatures.

container, and only a few of them are kicked up occasionally (Fig. 20-10A); this represents the model of a liquid with a few molecules evaporating from its surface. If we increase considerably the rotation speed of the wheel, all the Ping Pong balls will be thrown up into the air, and our device will now represent a model of a vapor or a gas. If the upper part of the "aquarium" were open, the fastest moving balls would get out of the container, and it would soon become quite empty. However, in the arrangement shown in Fig. 20-10, the container is provided with a movable piston which prevents the Ping Pong balls from leaving their enclosure. The balls that hit the piston and are reflected back into the container communicate to the piston a certain amount of mechanical momentum, and the total effect can be described as the gas pressure acting on the piston. Figure 20-10B, C, and D represents the situation that arises when there are various degrees of agitation of the molecules in our Ping Pong gas. In B, the velocities of the balls are not high enough to overcome the weight of the piston, so the piston hangs in the lowest position permitted by its suspension. In C, the balls move faster and push the piston up to a level position. In D, the velocity of the balls is still higher, and the piston is pushed all the way up.

20-5 Surface Tension and Surface Energy

We all are familiar with dewdrops and raindrops, and most of us have seen an escaped bit of mercury gather itself up into elusive droplets. Liquids all show this tendency to assume a characteristic spherical shape which competes with the force of gravity that forces liquids to assume the shape of their containers. A glass can be filled heaping full of water, so that the water level stands slightly above the rim and slopes down toward it at its edge.

These phenomena are examples of *surface tension*, apparent in all liquids. Surface tension can be explained from the molecular point of view by considering that each molecule is subject to attractive forces from its surrounding fellows. Let us consider this from the standpoint of potential energy. As always, we must first agree on what we shall consider to be the level of zero potential energy. As a matter of convenience, let us call the energy zero when a molecule is outside of the liquid we shall be considering. Now let us reach inside a drop of liquid with an imaginary pair of impossibly fine tweezers and remove one of the molecules. It is completely surrounded on all sides by other liquid molecules, and to remove it we shall have to do work against the attraction of all these neighbors. Since we must do this work to raise its energy to zero, it apparently had a quite large *negative* energy when it was inside the drop.

Having done this, let us repeat the process for a molecule on the outer surface of the drop. This molecule, being on the surface, is *not* surrounded by neighbors on all sides, but only on the bottom. Therefore, less work will have to be done to remove this molecule and bring it up to our agreed-on zero level of energy. And therefore, of course, its energy is *less negative* than the energy of an interior molecule.

So, because it is less negative, we see that a surface molecule has a greater potential energy than an interior molecule. (If you owe \$50, you are financially better off than if you owe \$100.) We have seen on many other occasions that physical systems come to a state of least energy if they are able to. (This is really a consequence of equipartition of energy and the second law of thermodynamics, although we shall not pursue the reasons any farther.) A truck rolls downhill, a hot body becomes cooler, and so on. And, to reduce the number of its high-energy surface molecules and therefore its total energy, a piece of liquid will do its best (against the pull of gravity) to make its surface as small as possible. The geometrical shape that has the least surface for a given volume is a sphere, so the liquid comes as close as it can to assuming the form of spherical drops.

20-6 Evaporation

The ideas of kinetic theory and of bonds between the molecules of a liquid help to explain the phenomenon of evaporation. We all know that water placed in a pan on a shelf will gradually disappear into the air; if the pan is put on a warm stove, it will evaporate much more rapidly. We must recall that whenever two bodies attract each other, it requires energy to separate them, whether the bodies are a heavy weight and the earth, or a pair of water molecules. If a water molecule happens to be in the top layer of molecules and happens to be moving upward and happens also to have speed enough to break away from the attractive

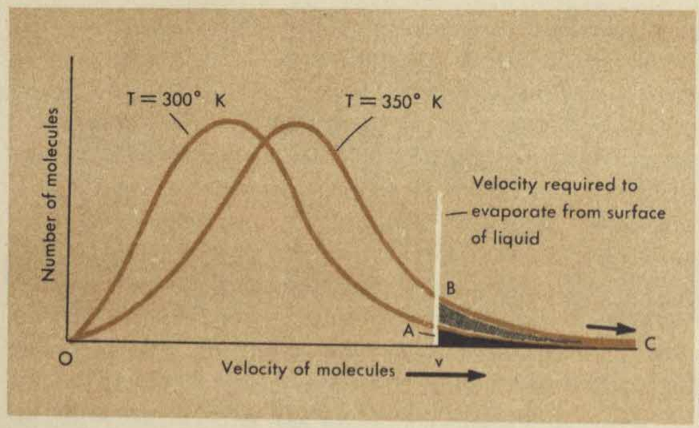


FIG. 20-11 The distribution of velocities in liquids of different temperatures; a slight rise in temperature gives many more molecules the energy needed to evaporate.

forces of its neighbors, it will then evaporate from the liquid surface and become a water-vapor molecule. Figure 20-11 shows the application to this situation of the Maxwellian distribution of molecular velocities, as was illustrated in Fig. 20-7. The velocity v is required for a water molecule to be able to pull free from the attraction of its fellows. At 300°K, only those molecules under the line AC, on the tail of the distribution curve, will have a speed of v or greater. This is a very small fraction of the total area under the 300° curve, and the evaporation rate will be slow. At 350°K, however, the curve is shifted to the right, and at this higher temperature all the molecules under BC will have enough speed to escape. This is a much greater fraction of the total number of molecules, and, as a consequence, the rate of evaporation will be increased a great deal more than would be expected from a mere comparison of temperatures.

We can also see from this molecular explanation of evaporation the reason why evaporation causes a liquid to become cooler. Only the fastest-moving and most energetic molecules are able to escape. With these high-energy molecules constantly leaving the liquid, the average energy of those remaining constantly becomes smaller, as long as the evaporation continues. Since temperature is really only a measure of the average kinetic energy of the molecules of a body, the evaporating liquid becomes cooler.

20-7 Monomolecular Layers

If matter were absolutely continuous, in principle we would be able to roll it into sheets of any desired small thickness. However, the molecular structure of matter places a lower limit on the possible thickness of material sheets or layers; they cannot be thinner than the diameter of

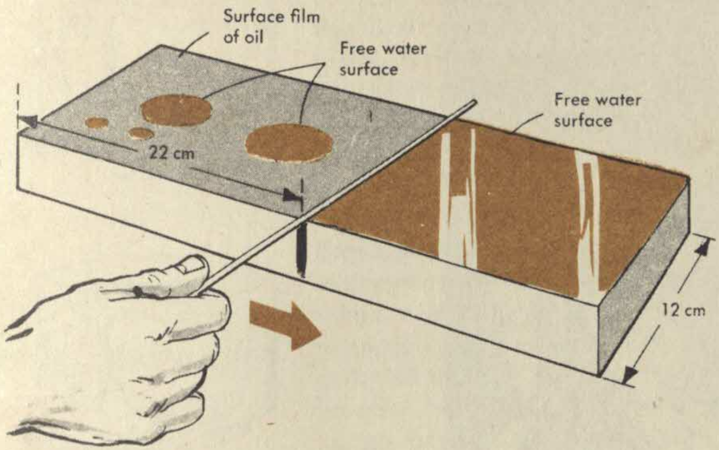


FIG. 20-12 Measuring the diameter of molecules by forming a monomolecular layer (a layer one molecule thick).

one single molecule. The most convenient way to produce such layers is to spread over the surface of water a small quantity of oil or some other substance not soluble in water. Let us take a long troughlike container filled to the brim and place a wire across it as indicated in Fig. 20-12. For a trough of reasonable size, it is difficult to measure directly a small enough amount of oily substance to cover only a part of its surface. So let us take a gram of stearic acid (one of the constituents of many fats and waxes) and dissolve it in 1000 cm^3 of the very volatile solvent acetone. Before we put a drop of this solution on the water surface, we must find the volume of a drop. This is easily done by measuring the volume of, say, 100 drops from the dropper we are going to use. Assume this volume to be 2.3 cm^3, so that the volume of a single drop is 2.3×10^{-2} cm^3. Now, how much stearic acid does this single drop contain? A handbook gives the information that stearic acid has a density of 0.94 gm/cm^3; the original gram that we dissolved thus had a volume of $1/0.94 = 1.06$ cm^3. We dissolved this in 1000 cm^3 of solvent, so that each cm^3 of the solution contained 1.06×10^{-3} cm^3 of stearic acid. In the single drop, then, there is $1.06 \times 10^{-3} \times 2.3 \times 10^{-2} = 2.4 \times 10^{-5}$ cm^3 of stearic acid.

We now drop a single drop on the surface of the water-filled trough, which is 12 cm wide. In a few seconds, the acetone solvent evaporates, leaving only 2.4×10^{-5} cm^3 of stearic acid to form a film on the surface. Now, the crosswire is moved slowly and carefully to the right (in Fig. 20-12) to stretch the film as much as we can. We find that the wire cannot be moved beyond 22 cm from the end of the trough without patches of bare water surface being formed. We thus conclude that $12 \times 22 = 264$ cm^2 is the area of the monomolecular film formed

FIG. 20-13 A "random walk," in which a person (or other particle) frequently changes his direction by turning at unpredictable angles.

by 2.4×10^{-5} cm^3 of stearic acid. The thickness of the film is the volume divided by the area, or $2.4 \times 10^{-5}/264 = 9 \times 10^{-8}$ cm. We may thus from this rough experiment conclude that the diameter of a molecule of stearic acid is in the neighborhood of 10^{-7} cm. This compares quite reasonably with the approximate value of 4×10^{-8} cm which we previously calculated as the diameter of a molecule of air. We would rightly expect the stearic acid molecule to be larger than an air molecule, so our figures, rough as they are, seem to be quite consistent.

20-8 Diffusion

The phenomenon of *diffusion* takes place in liquids as well as in gases. If we drop a lump of sugar into a cup of tea and then do not stir it, it

will take a very long time, many hours as a matter of fact, until the tea becomes uniformly sweetened. If we *do* wait, we shall find that the situation is the same as could have been achieved in a minute by using a spoon. How do the molecules of sugar, without being stirred, travel through the liquid? Well, they just shoulder their way through the crowd of water molecules in the cup as the result of thermal agitation. They do not follow any straight path but execute a "random walk," as illustrated in Fig. 20-13.

It is worth turning back to Fig. 20-3 to note that the motion of a Brownian particle is also a "random walk." The same arguments apply to a smoke particle in collision with air molecules as to a sugar molecule under bombardment by water molecules. If a person takes N steps all in one direction and the length of each step is L, then obviously he will cover the distance NL away from the point from which he started. If, however, he changes his direction at random with each step, he does not go that far. It can be shown mathematically in this case that the mean distance traveled away from the original point, averaged over many trials, is only $\sqrt{N} \times L$. Thus, since the number of steps taken is proportional to time, the distance traveled in a random walk increases only as the square root of time. The situation is illustrated in Fig. 20-14, which shows the diffusion of sugar molecules through the unstirred cup of tea. The phenomenon of diffusion plays an important role in many other processes of physics besides the sweetening of unstirred tea. The principle of the atomic reactor is based on the diffusion of neutrons through the moderator, and the energy quanta produced by thermonuclear reactions in the center of the sun similarly diffuse through the body of the sun until they can fly off freely into space after reaching the surface.

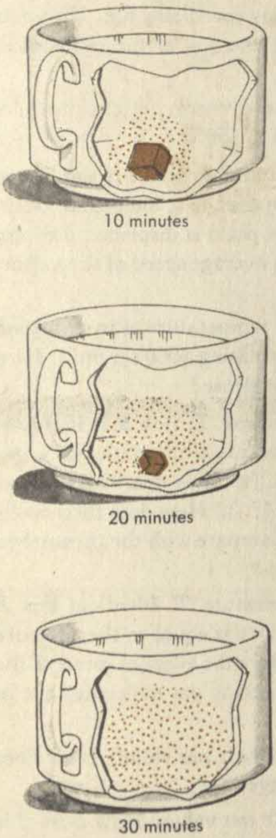


FIG. 20-14 The diffusion of sugar in an unstirred teacup. The distances diffused in 10, 20, and 30 minutes are in the ratio of $\sqrt{1} : \sqrt{2} : \sqrt{3}$, or 1 : 1.41 : 1.73.

Questions

(20-1)

1. Liquid sulfur dioxide has a density of 1.4 gm/cm^3. In the gaseous state under normal conditions, its density is 2.9×10^{-3} gm/cm^3. By about how many molecular diameters are the gaseous molecules separated?

2. The element bromine is a liquid at room temperature, with a density of 3.12 gm/cm^3. When vaporized (at 60°C) its density is 7.59×10^{-3} gm/cm^3. By about how many molecular diameters are its gaseous molecules separated?

(20-2)

3. The text gives 0.65 cm/sec as the average velocity of particles 1 micron in diameter at 27°C. What would be the average speed of particles 10 times this large in diameter, made of the same material, and at the same temperature?

4. Observation shows that at the same 27°C certain uniform suspended particles have an average speed of 1.30 cm/sec, rather than the 0.65 cm/sec quoted in the text. What is the diameter of these faster-moving particles? (Assume them to have the same density.)

5. Molecules of gas at $-73°C$ have a certain average speed and KE. To what temperature must the gas be heated to (a) double the KE of the molecules; (b) double the average speed of the molecules?

6. By what factor are (a) the velocity; (b) the kinetic energy, of the particles in Question 3 multiplied if the temperature is raised to 159°C?

(20-3)

7. A molecular-beam apparatus such as described in the text has a drum 10 cm in diameter which revolves 5400 rev/min. When exposed to a molecular beam of vaporized sodium, the most dense deposit on the plate is displaced 0.60 cm from the point opposite the entrance slit. What is the average speed of the vaporized sodium atoms?

8. Consider an oven containing sodium at the same temperature as in Question 7, with a similar rotating drum 12 cm in diameter, rotating 4000 rev/min. How far from point A will the deposit of sodium be most dense?

(20-4)

9. Box A contains N molecules of a gas at temperature T. Box B is identical with A and is at the same temperature; but B contains N molecules of a gas whose molecules are twice as massive as those in A. (a) How does the average energy of the molecules in B compare with those in A? (b) How does their average speed compare? (c) How does the pressure in B compare with the pressure in A?

10. Box A contains N molecules of a gas at temperature T. Identical box B contains some of the same gas that is in A; its pressure is equal to the pressure in A, but its absolute temperature is $1.5\,T$. (a) How does the average speed of the molecules in B compare with those in A? (b) How many gas molecules are in box B?

11. Take the numbers 1, 2, 3, 4, 5, and 6. What is their *rms* value? How does this compare with their simple arithmetical average?

12. Take the numbers 2, 4, 6, and 8. What is their *rms* value? How does this compare with their arithmetical average?

13. A molecule of hydrogen has a mass $\frac{1}{2}$ that of a molecule of helium. How do their *rms* speeds compare, if both gases are at the same temperature?

14. A krypton molecule is 2.25 times as massive as a molecule of oxygen. If samples of both gases are at the same temperature, how do their *rms* speeds compare?

15. The density of helium gas at 1 atm. pressure and 0°C is 1.78×10^{-4} gm/cm^3. What is the *rms* speed of the helium molecules?

16. The density of chlorine gas is 3.2×10^{-3} gm/cm^3 at 1 atm. pressure and 0°C. What is the *rms* speed of chlorine molecules at this temperature?

(20-5)

17. Consider a single droplet of 1 gm of mercury on a level surface. Now, with a knife or card, divide the droplet into two 0.5-gm droplets. (a) Does the mercury now have more or less surface area than it did as a single droplet? (b) Does it now have more or fewer mercury atoms on the surface? (c) Has its surface energy been increased or decreased?

18. (See Question 17.) Explain why two clean mercury droplets, when pushed into contact, will spontaneously run together to form a single droplet. Will the single droplet formed in this way be a little bit warmer, or a little bit cooler, than the separate pair of droplets?

(20-6) **19.** A pan of water at 27°C is set in the breeze on a warm (27°C) dry day. The water cools to 22°C, after which its temperature does not change. Explain why it does not continue to cool until it is all evaporated away.

20. If a pan of ether (or alcohol, or other relatively volatile liquid) at 27°C is set in the same breeze on the same day as the pan of water in Question 19, will it become cooler than the pan of water? Explain.

chapter / twenty-one

The Electrical Nature of Matter

21-1 Positive and Negative Ions

Pure distilled water is a very poor conductor of electricity. However, if we dissolve in water a small amount of some acid or base or salt, its electrical conductivity becomes quite appreciable. In contrast to the case of metallic conductors, the passage of electric current through water solutions is associated with certain chemical phenomena, the nature of which depends on the particular solute used. If we pass an electric current through a solution of nitric acid (HNO_3), small gas bubbles will be formed on both electrodes and will gradually rise to the surface. We can collect these gases in two long inverted glass cylinders that are placed about the electrodes and that are originally completely filled with water (Fig. 21-1A). When we analyze the nature of the gas liberated on the negative electrode, we find it to be hydrogen; in fact, if we open the valve at the top of the glass cylinder placed above this electrode, we can ignite the gas streaming out from it, and in the process of burning, the hydrogen will unite with atmospheric oxygen and form water vapor. The gas that is collected in the cylinder placed above the positive

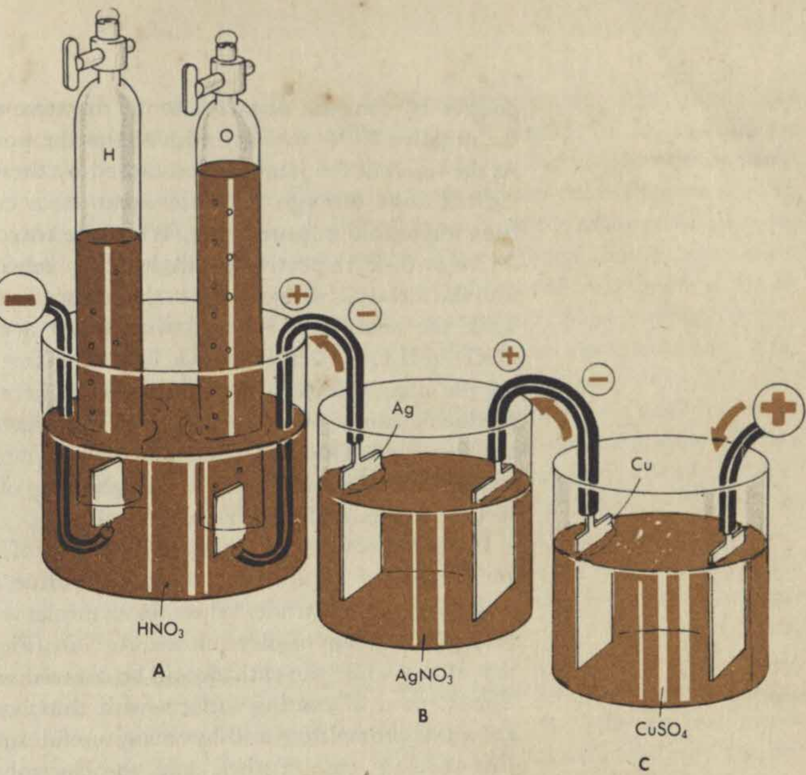


FIG. 21-1 The electrolysis of water solutions of (A) nitric acid, (B) silver nitrate, and (C) copper sulfate by the same current which passes through the three cells in series.

electrode is oxygen; if we open the valve at the top of that cylinder and place a burning match into the stream of outcoming gas, it will flare up more intensely because of the additional oxygen supply.

Thus the passage of electric current decomposes water into its two elementary constituents, hydrogen and oxygen. How does this happen, and why should it require something dissolved in the water to get things going? The water molecule H_2O is quite tightly bound together and has very little tendency to break apart into ions. In fact, at room temperature, only about 1 molecule in 10^7 will be split up into H^+ and OH^- ions. Such a small number of ions migrating through an electric field constitutes so small a current that pure distilled water may be considered a fairly good insulator. However, the nitric acid molecules (as well as the molecules of salts and bases) split up readily into H^+ and NO_3^- when dissolved in water. Thus a large number of charged ions are provided, and their migration in the electric field between anode and cathode can constitute a large current.

When an electric potential is applied to the cathode and anode, the

positive H^+ ions are attracted to the negative electrode (cathode) and the negative NO_3^- ions are attracted to the positive electrode (anode). As the result of the ionic motion caused by these attractions, an electric current flows through the water solution, a current that would have been impossible in pure water. When the traveling ions H^+ and NO_3^- arrive at their respective electrodes, they release their electric charges into the metal; hydrogen rises to the surface in the form of tiny bubbles, while the neutral NO_3 reacts with water according to the equation $2NO_3 + H_2O \rightarrow 2HNO_3 + O$, liberating free oxygen and regenerating the original molecules of nitric acid. (There are secondary reactions producing oxides of nitrogen, but we may leave these complications to the chemists.) Thus the passage of electric current through the water resulted in nothing more than the breaking up of the water molecules into their hydrogen and oxygen components.

If, instead of using nitric acid, we use one of its salts (in which hydrogen is replaced by a metal), the metal will be deposited on the surface of the negative electrode. When, for example, we pass an electric current through a solution of silver nitrate, $AgNO_3$ (Fig. 21-1B), we shall notice that after a while the cathode will be covered with a thin layer of silver. This method of coating surfaces with thin layers of various metals is known as *electroplating* and has many useful and practical applications. Just as in the case of nitric acid, the electrolytic process in the silver nitrate solution is due to the fact that the molecules of this salt break up into two oppositely charged ions Ag^+ and NO_3^-, which are driven in opposite directions by the applied electric potential. When the Ag^+ ions reach the negative cathode, they pick up their lost electrons from the cathode and become neutral insoluble Ag atoms which form the cathode deposit. At the anode, the NO_3^- ions give up their electrons to the anode, so the net effect is that of a stream of electrons flowing within the electrolysis tank from cathode to anode.

21-2 The Laws of Faraday

Michael Faraday, whose name has already been mentioned in connection with the theory of electric and magnetic fields, was the first to investigate in detail the laws of electrolytic processes. He found first of all that, for each given electrolyte solution, the amount of material deposited at the electrodes is directly proportional to the strength of the electric current and to its duration, or, in other words, that ***the amount of material deposited on the electrodes is directly proportional to the total amount of electric charge which passes through the solution.*** From this *first law of Faraday*, we conclude that each ion of a given chemical substance carries a well-defined electric charge.

In further studies, Faraday investigated the relative amounts of electric charge carried by ions of several different chemical substances. To compare these amounts, he passed an electric current consecutively

through the solutions of such substances as nitric acid, silver nitrate, and copper sulfate, as is shown in Fig. 21-1A, B, and C. In the case of nitric acid, a certain amount of hydrogen gas was liberated on the cathode, while a certain amount of silver was deposited on the cathode in the case of the silver nitrate solution. Faraday measured the amounts of hydrogen and silver produced in these experiments and found that the ratio of the weight of deposited silver to the weight of liberated hydrogen was 107.02.

Chemists had before this time determined from many ingenious experimental measurements the *relative* weights of the atoms of many of the chemical elements. So long as we are dealing with relative weights—the ratio of the mass of one atom to the mass of another—it really makes no difference what we use as a standard. If the ratio of the mass of a hydrogen atom to the mass of an oxygen atom is about 1/16, it is also 2/32, or 1.96/31.36. In the early days of analytical chemistry, hydrogen (the lightest of the elements) was arbitrarily assigned the value of exactly 1. On this scale, the mass of the oxygen atom would be 15.87. Later (primarily because most of the analyses to determine atomic weights were done with oxides of the elements), it was decided to make oxygen exactly 16, which required that hydrogen be raised to 1.008. In recent years, there has been another change. As we shall see later, not all oxygen atoms are alike, nor are all carbon atoms. In fact, the atoms of nearly all the elements have several different *isotopes*, which differ from one another only in their masses. The present standard for both physicists and chemists is the C-12 isotope of carbon, which is assigned the value of exactly 12. In other words, the unit of mass in determining the relative masses (or "weights," in common usage) is 1/12 of the mass of one atom of C-12. Table 21-1 gives the relative masses of the 103 known elements relative to this standard. From this table, we can see that 107.02 is exactly the ratio of the atomic weight of silver to the atomic weight of hydrogen. Thus Faraday concluded that the same number of atoms of Ag and H had been deposited and that *one ion of silver carries exactly the same electric charge as one ion of hydrogen.* It would be premature, however, to conclude that *all* ions carry the same electric charge. In fact, comparing the amount of silver liberated in the electrolysis of silver nitrate with that of copper liberated by the same electric current flowing for the same length of time in the electrolysis of copper sulfate, we find that the weight ratio of silver to copper is 3.40 instead of the 1.70 (i.e., 107.9/63.5) that would correspond to 1 atom of silver per atom of copper. Notice, however, that 1.70 is exactly one-half of 3.40, and if we write the observed ratio in the form $2 \times 107.9/63.5$, we conclude that *one ion of copper carries twice as much electricity as one ion of silver.*

We can interpret this by saying that the copper ion has exactly twice the electric charge of the silver ion. We know, as the nineteenth-century scientists did not, that this is because the silver ion has lost one electron, while the copper ion has lost two electrons and therefore has a double

positive charge. The number of electrons lost or gained by an ion is one aspect of what the chemist calls *valence*. Thus hydrogen, silver, and the nitrate group (NO_3) have a valence of 1 (are *monovalent*), copper and the sulfate group (SO_4) have a valence of 2 (are *divalent*), while aluminum ions have a valence of 3 (are *trivalent*).

Thus with several electrolytic cells in series, as in Fig. 21-1, for eac atom of a monovalent element that is deposited, only $\frac{1}{2}$ of a divalent atom, or $\frac{1}{3}$ of a trivalent atom can be deposited. Chemists call the atomic weight divided by the valence the *equivalent weight*, and *Faraday's second law of electrolysis* states that ***when the same amount of electric charge flows through different electrolytic cells, the amounts of the substances deposited (or liberated) are in direct proportion to their equivalent weights.***

For example, we can place two cells in series (which guarantees that the same amount of charge will flow through each), one cell containing silver nitrate and the other gold chloride (gold is trivalent), and allow current to flow until we have 1.00 gm of silver deposited on the cathode of the first cell. At this time, how much gold will have been deposited on the cathode of the other cell? The equivalent weight of silver, since silver is monovalent, is the same as its atomic weight, or 107.9. Gold has an equivalent weight of $197.0/3 = 65.7$. Therefore, we can write

$$\frac{\text{weight Ag deposited}}{\text{weight Au deposited}} = \frac{1.00}{x} = \frac{107.9}{65.7}$$

or

$$x = 0.609 \text{ gm Au deposited.}$$

It has been found that the passage of 96,500 coulombs of charge (this amount of charge is known as a *faraday*) will deposit a mass, in grams, of any element which is numerically equal to its equivalent weight. (This amount of any element is more formally called a *gram-equivalent weight*; one *gram-atomic weight* is, of course, an amount of substance whose mass in grams equals its atomic weight.) The American physicist Robert A. Millikan (1868–1953) showed with his famous oil-drop experiment (which we shall discuss later in this chapter) that the value of a single electronic charge is 1.60×10^{-19} coul or 4.80×10^{-10} esu. Thus the passage of 96,500 coul means the passage from cathode to anode of $96{,}500/(1.60 \times 10^{-19}) = 6.02 \times 10^{23}$ electrons. These electronic charges are escorted across one at a time by monovalent ions, two at a time by divalent ions, and so on. This leads us to the important conclusion that ***a gram-atomic weight of any element contains 6.02×10^{23} atoms.*** The number, 6.02×10^{23}, is known as *Avogadro's number*, after the nineteenth-century Italian chemist who first reasoned that gram-molecular weights (the extension of gram-atomic weight to gram-molecular weight should be obvious) of all gases contained the same number of molecules.

TABLE 21-1 LIST OF CHEMICAL ELEMENTS

Name	Symbol	At. No.	Description	At. Wt.	Name	Symbol	At. No.	Description	At. Wt.
*Actinium**	Ac	89		(227)‡	[*Mendelevium*]	Mv	101		(256)
Aluminum	Al	13	White metal	26.98	Mercury	Hg	80	White liquid metal	200.59
[*Americium*]†	AM	95	White metal	(243)	Molybdenum	Mo	42	White metal	95.94
Antimony	Sb	51	Gray solid	121.76	Neodymium	Nd	60	Yellow metal	144.24
Argon	A	18	Colorless gas	39.94	Neon	Ne	10	Colorless gas	20.18
Arsenic	As	33	Gray solid	74.92	[*Neptunium*]	Np	93	White metal	(237)
[*Astatine*]	At	85		(210)	Nickel	Ni	28	White metal	58.37
Barium	Ba	56	White metal	137.34	Niobium	Nb	41	Gray metal	92.91
[*Berkelium*]	Bk	97		(249)	Nitrogen	N	7	Colorless gas	14.01
Beryllium	Be	4	White metal	9.01	[*Nobelium*]	No	102		(254)
Bismuth	Bi	83	White metal	208.98	Osmium	Os	76	White metal	190.2
Boron	B	5	Black solid	10.81	Oxygen	O	8	Colorless gas	16.00
Bromine	Br	35	Brown liquid	79.91	Palladium	Pd	46	White metal	106.4
Cadmium	Cd	48	White metal	112.40	Phosphorus	P	15	Red or white solid	30.97
Calcium	Ca	20	White metal	40.08	Platinum	Pt	78	White metal	195.09
[*Californium*]	Cf	98		(251)	[*Plutonium*]	Pu	94	Metal	(242)
Carbon	C	6	Transparent crystal or black solid	12.01	*Polonium*	Po	84	Metal	(210)
Cerium	Ce	58	Gray metal	140.12	Potassium	K	19	White metal	39.10
Cesium	Cs	55	White metal	132.91	Praseodymium	Pr	59	Yellow metal	140.91
Chlorine	Cl	17	Greenish gas	35.45	[*Promethium*]	Pm	61		(145)
Chromium	Cr	24	White metal	52.00	*Protactinium*	Pa	91	White metal	(231)
Cobalt	Co	27	Gray metal	58.93	*Radium*	Ra	88	White metal	(226)
Copper	Cu	29	Reddish metal	63.54	*Radon*	Rn	86	Colorless gas	(222)
[*Curium*]	Cm	96	White metal	(247)	Rhenium	Re	75	Metal	186.22
Dysprosium	Dy	66	Gray metal	162.50	Rhodium	Rh	45	White metal	102.91
[*Einsteinium*]	E	99		(25.4)	Rubidium	Rb	37	White metal	85.47
Erbium	Er	68	Gray metal	167.26	Ruthenium	Ru	44	Gray metal	101.07
Europium	Eu	63	Gray metal	151.96	Samarium	Sm	62	Gray metal	150.35
[*Fermium*]	Fm	100		(25.3)	Scandium	Sc	21	Metal	44.96
Fluorine	F	9	Yellowish gas	19.00	Selenium	Se	34	Red or gray solid	78.96
[*Francium*]	Fr	87		(223)	Silicon	Si	14	Gray metalloid	28.09
Gadolinium	Gd	64	Gray metal	157.25	Silver	Ag	47	White metal	107.87
Gallium	Ga	31	Gray solid	69.72	Sodium	Na	11	White metal	22.99
Germanium	Ge	32	Gray metal	72.59	Strontium	Sr	38	White metal	87.62
Gold	Au	79	Yellow metal	196.97	Sulfur	S	16	Yellow solid	32.06
Hafnium	Hf	72	White metal	178.49	Tantalum	Ta	73	Metal	180.95
Helium	He	2	Colorless gas	4.003	[*Technetium*]	Tc	43	Gray metal	(99)
Holmium	Ho	67	Gray metal	164.93	Tellurium	Te	52	Gray metal	127.60
Hydrogen	H	1	Colorless gas	1.008	Terbium	Tb	65	Metal	158.92
Indium	In	49	White metal	114.82	Thallium	Tl	81	Gray metal	204.37
Iodine	I	53	Black solid	126.90	*Thorium*	Th	90	White metal	232.04
Iridium	Ir	77	White metal	192.2	Thulium	Tm	69	Metal	168.93
Iron	Fe	26	White metal	55.85	Tin	Sn	50	White metal	118.69
Krypton	Kr	36	Colorless gas	83.80	Titanium	Ti	22	White metal	47.90
Lanthanum	La	57	White metal	138.91	Tungsten	W	74	Metal	183.85
[*Lawrencium*]	Lw	103		(257)	*Uranium*	U	92	White metal	238.03
Lead	Pb	82	Gray metal	207.19	Vanadium	V	23	White metal	50.94
Lithium	Li	3	White metal	6.94	Xenon	Xe	54	Colorless gas	131.30
Lutetium	Lu	71	Gray metal	174.97	Ytterbium	Yb	70	Metal	173.04
Magnesium	Mg	12	Gray metal	24.31	Yttrium	Y	39	Metal	88.91
Manganese	Mn	25	Gray metal	54.94	Zinc	Zn	30	White metal	65.37
					Zirconium	Zr	40	Metal	91.22

* Italic type indicates radioactive element.

† Square brackets indicate element does not occur in nature in more than, perhaps, insignificant traces.

‡ Parentheses indicate mass number of longest-lived isotope.

FIG. 21-2 Sir J. J. Thomson (left), discoverer of the electron, and Lord Rutherford, discoverer of the nucleus, discuss problems in the courtyard of the Cavendish Laboratory, Cambridge, England, 1929.

21-3 The Passage of Electricity through Gases

The next step in the study of the electric nature of matter was made by another famous Britisher named J. J. Thomson (1856–1940) (Fig. 21-2). Whereas Faraday studied the passage of electric current through liquids, J. J. (as he was known to his colleagues and his students) later concentrated his attention on the electrical conductivity of gases.

When we walk in the evening along the downtown streets of a modern city, we observe the bright display of neon (bright red) and helium (pale green) advertising signs. Modern offices and homes are illuminated by fluorescent light tubes. In all these cases, we deal with the passage of high-voltage electric current through a rarefied gas—the phenomenon that was the object of the lifelong studies of J. J. Thomson. As in the case of liquids, the current passing through a gas is due to the motion of positive and negative ions driven in opposite directions by an applied electric field. The positive gas ions are similar to those encountered in the electrolysis of liquids (being the positively charged atoms or molecules of the substance in question), and the negative ions in this case are the much less massive singly charged particles that we now know to be electrons.

To study these particles, mysterious at that time, Thomson, in 1897, used an instrument shown schematically in Fig. 21-3. It consisted of a glass tube containing highly rarefied gas with a cathode placed at one end

of it and an anode located in an extension on the side. Because of this arrangement, the negative ions, which form the "cathode rays" that move from left to right in the drawing, miss the anode and fly into the right side of the tube. The tube broadens here, and its flat rear end is covered with a layer of fluorescent material which becomes luminous when bombarded by fast-moving particles. This tube is very similar to a modern TV tube, in which the image is also due to the fluorescence produced by a scanning electron beam. But, in those pioneering days of what we now call "electronics," one was satisfied with much simpler shows; placing a metal cross in the way of the beam, Thomson observed that it cast a shadow on the fluorescent screen, indicating that the particles in question were moving along straight lines, similar to light rays.

21-4 The Charge-to-mass Ratio of an Electron

Thomson's next task was to study the deflection of the beam caused by electric and magnetic fields applied along its path. Indeed, since the beam was formed by a swarm of negatively charged particles, it should be deflected toward the positive one of the parallel plates that produce the electric field shown in Fig. 21-4A. On the other hand, the beam of charged particles should be deflected by a magnetic field directed perpendicularly to its track (Fig. 21-4B) according to the laws of electromagnetic interactions.

The deflection of a particle will depend, of course, on how much force is applied to it. For a charged particle in an electric field, the force

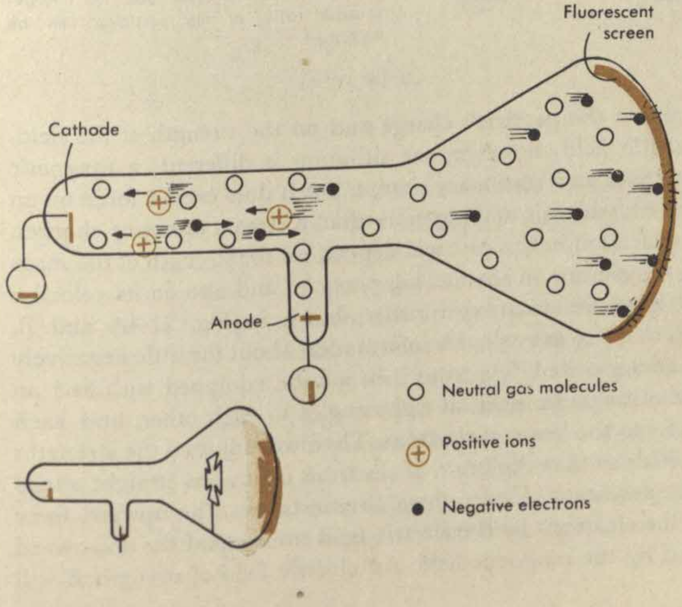


FIG. 21-3 The passage of electric current, carried by both positive and negative particles through a rarefied gas.

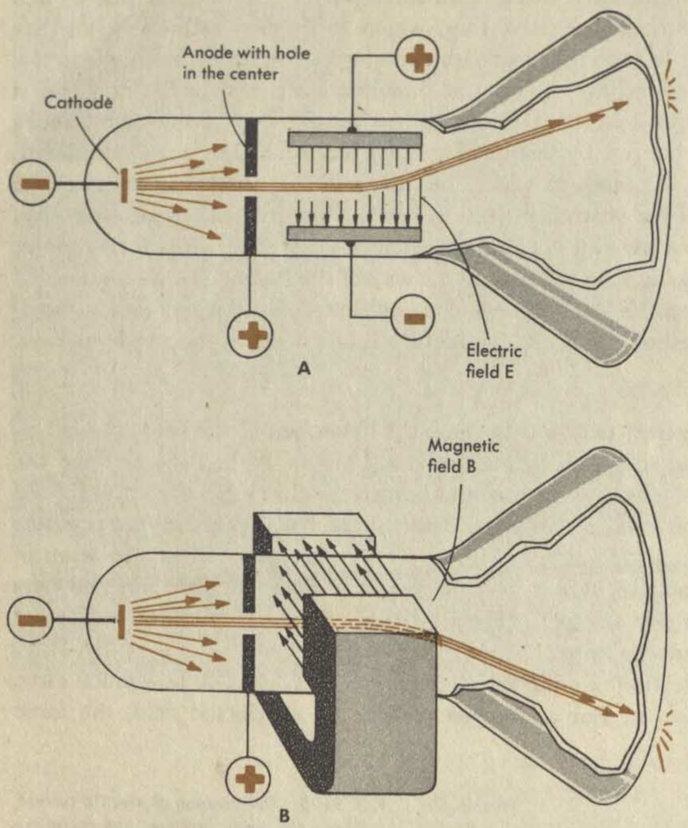


FIG. 21-4 (A) the deflection of negative particles (electrons) by an electric field and (B) by a magnetic field. By combining both deflections in the same tube, the charge-to-mass ratio of the particles can be measured.

depends only on the particle's charge and on the strength of the field. For a magnetic field, however, the situation is different; a magnetic field has no effect on a stationary charge, but it does exert a force on an electric current, which is nothing more than a stream of *moving* charges. Hence the deflection in this case will depend on the strength of the magnetic field, the charge on the moving particle, and also on its velocity.

By combining the two experiments shown in Fig. 21-4A and B, Thomson was able to get valuable information about the little negatively charged particles called "electrons." In a tube equipped with *both* an electric and a magnetic field, at right angles to each other, and each at right angles to the beam of electrons, Thomson adjusted the strengths of the two fields so that the beam of electrons continued straight ahead without any deviation. Under these circumstances, the upward force exerted on the electrons by the electric field must equal the downward force exerted by the magnetic field. An electric field of strength E will

exert a force of E newtons on a charge of 1 coulomb; on an electron whose charge is e, the force will be Ee. In the magnetic field, we have seen that the force on a moving charge is Bvq, which for our electron becomes Bve. If we set these two forces equal, we get

$$Ee = Bve$$

or

$$v = \frac{E}{B}.$$

This equation does not even include the mass or the charge of the electron but does serve to measure the velocity the electrons have if they proceed undeviated by the given B and E. The experimenter knows, however, the potential difference between the anode and cathode of his tube, which we can call V volts. By the definition of a volt, a charge of 1 coul accelerated from cathode to anode would have had V joules of work done on it and hence would pass through the hole in the anode with an equal amount of kinetic energy V. For an electron of charge e coul, the energy would be Ve, which we can set down as equal to the $\frac{1}{2}mv^2$ of the electron:

$$Ve = \tfrac{1}{2}\, mv^2.$$

In this equation, it is important to recall that if V is in volts (joules per coulomb) and e is in coulombs, the energy will be expressed in joules. The mass of the electron m must therefore be in kilograms, and its velocity v in meters per second. When we substitute into this equation the previously determined value of v, we get

$$Ve = \frac{mE^2}{2B^2}$$

from which

$$\frac{e}{m} = \frac{E^2}{2VB^2}.$$

Thus we can measure the charge-to-mass ratio of the electron in terms of the readily measurable quantities E, B, and V that are used in the tube.

Actually, Thomson used a somewhat different method, which was based on measuring the deflection of the electron beam by the electric field alone when the magnetic field was turned off. This leads into complications involving the geometry of the tube that we can avoid by using this simpler method of calculation, which provides equivalent results. Thomson's experiments, and those of later workers, give the value $e/m = 1.76 \times 10^{11}$ coul/kg, or 1.76×10^{8} coul/gm. Unfortunately, he was not then able to solve his equations to determine e, the charge of the electron, because at that time the mass of the electron was not known.

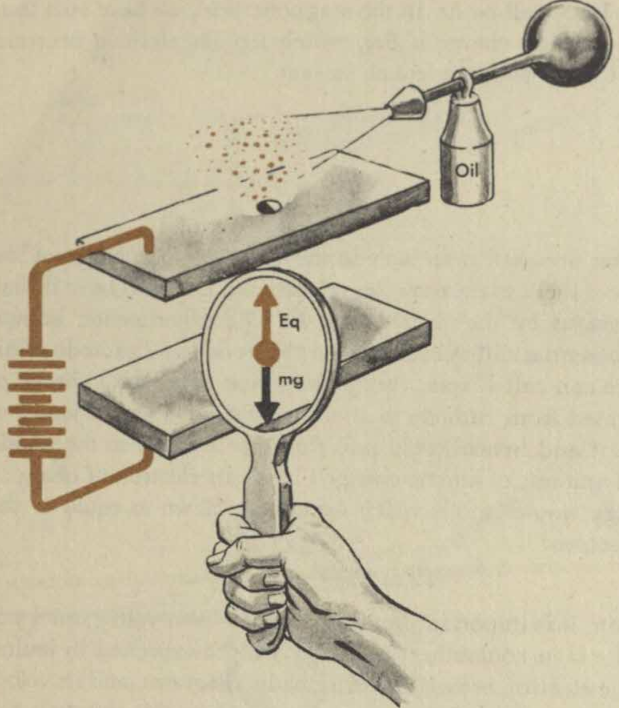


FIG. 21-5 A schematic diagram of Millikan's experiment for measuring the elementary electron charge.

21-5 The Charge and Mass of an Electron

This work of Thomson's paved the way for the work of the celebrated American physicist Robert A. Millikan, who directly measured the charge of the electron by means of a very ingenious experiment illustrated in Fig. 21-5. A cloud of tiny oil droplets was sprayed into the space above the plates, and a small hole in the top plate was uncovered long enough for one of the droplets to drift down through the hole into the space between the plates, where it could be observed through a microscope set into the wall of the vessel. By means of a relationship known as Stokes' law, the weight of a small droplet can be determined from the rate at which it settles downward through the air. Millikan was able to measure the rate of settling with no electric field between the plates and thus compute the weight of the droplet.

Ultraviolet light can drive electrons away from the molecules of objects on which it falls, so by allowing a beam of ultraviolet light to shine between the plates, Millikan was able to cause the droplet to have a slight charge that could change suddenly from time to time as it collided with charged air molecules. By varying the potential applied across

his plates, he could adjust the electric field until the droplet would hang motionless, neither rising nor falling. Under these equilibrium conditions, the upward force caused by the electric field was just equal to the weight of the droplet, and thus

$$Eq = mg$$

from which q, the charge on the droplet, could be easily figured.

It turned out that all the charges measured in this way were small *integral* multiples of a certain quantity that was apparently the elementary electric charge, or the charge of an electron. Numerically, he found that the value of this elementary charge is 1.60×10^{-19} coul, or 4.80×10^{-10} esu.

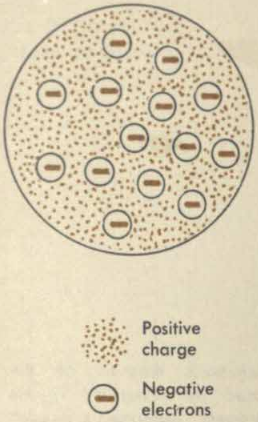


FIG. 21-6 J. J. Thomson's atomic model, showing negative electrons floating in a ball of positive "electric fluid."

From Thomson's charge-to-mass ratio and a direct knowledge of the charge on an electron, the mass of an electron can be computed to be

$$\frac{1.60 \times 10^{-19}\ \text{coul}}{1.76 \times 10^{8}\ \text{coul/gm}} = 9.11 \times 10^{-28}\ \text{gm}.$$

The discovery of the electron as representing a free electric charge and the possibility of its extraction from neutral atoms was the first indication that *atoms are not indivisible particles but complex mechanical systems composed of positively and negatively charged parts.* Positive ions were interpreted as having a *deficiency* of one or more electrons, whereas negative ions were considered as atoms having an *excess* of electrons.

21-6 Thomson's Atomic Model

On the basis of his experiments, J. J. Thomson proposed a model of internal atomic structure (Fig. 21-6) according to which atoms consisted of a positively charged substance (*positive electric fluid*) distributed uniformly over the entire body of the atom, with negative electrons imbedded in this continuous positive charge like seeds in a watermelon. Since electrons repel each other but are, on the other hand, attracted to the center of the positive charge, they were supposed to assume certain stable positions inside the body of the atom. If this distribution were disturbed by some external force, such as, for example, a violent collision between two atoms in a hot gas, the electrons were supposed to start vibrating around their equilibrium positions, emitting light waves of corresponding frequencies.

At this time (1904), it was well known that each element, when in the form of an incandescent vapor, would emit light of only certain frequencies. Many calculations were made in an attempt to reconcile the theoretical frequencies of vibrations that could be expected from Thomson's electrons with the actually observed frequencies of the light emitted by various elements. All these efforts were failures, and it became

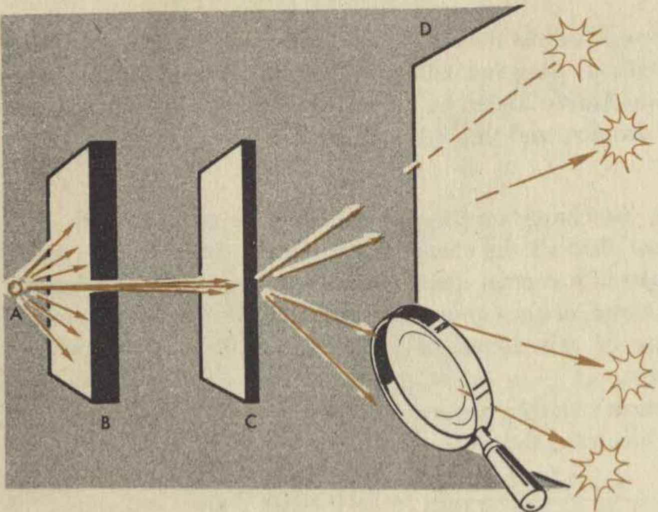


FIG. 21–7 Schematic diagram of the arrangement used by Rutherford in his "atomic bombardment" experiments through which he discovered the nucleus.

clear that there was something radically wrong with Thomson's model. It considered an atom to be a complex structure of positive and negative charge, and thus represented a considerable progress toward the truth; but it was not the true picture of an atom.

21-7 Rutherford's Atomic Model

The honor of giving the first correct description of the distribution of positive and negative charges within the atom belongs to a New Zealand-born physicist, Ernest Rutherford (1871–1937) (Fig. 21-2), who was later elevated to the rank of Lord Rutherford for his important scientific achievements. Young Rutherford entered physics during that crucial period of its development when the phenomenon of natural radioactivity had just been discovered, and he was the first to realize that radioactive phenomena represent a spontaneous disintegration of heavy unstable atoms.

Radioactive elements emit three different kinds of rays: high-frequency electromagnetic waves known as γ *rays* ("gamma rays"), beams of fast-moving electrons known as β *rays* ("beta rays"), and α *rays* ("alpha rays"), which were shown by Rutherford to be streams of very-fast-moving helium ions. Rutherford realized that important information about the inner structure of atoms could be obtained by the study of collisions between onrushing α particles and the atoms of various materials forming the target. This started him on a series of epoch-making atomic bombardment experiments that revealed the true nature of the atom. The basic idea of the experimental arrangement used by Rutherford in his studies was exceedingly simple (Fig. 21-7): a speck of α-

emitting radioactive material at A, a lead diaphragm B that cut out a thin beam of α rays, the material under investigation in the form of a piece of thin foil C, a fluorescent screen D, and a microscope E to observe the tiny flashes of light, or scintillations, originating when an α particle hits the screen.

How could one expect the α particles to be deflected according to Thomson's model of the atom? Alpha particles (their structure then unknown) were known to be doubly charged positive ions of helium; they acted as very efficient projectiles in being able to penetrate through thin metal foils at least several hundred atoms in thickness. Mathematical analysis showed that such a positive projectile, after penetrating several hundred of Thomson's spheres of positive charge, would be deflected by electrostatic forces, but the total deflection could not possibly add up to more than a few degrees.

But the results of the experiment, when performed by two of Rutherford's students, Hans Geiger and Ernest Marsden, were very different. Most of the α particles penetrated the foil with very little deflection. An appreciable fraction of them, however, were deflected through large angles—a few were turned back almost as though they had been reflected from the foil. This was a deflection of nearly 180°, and a completely impossible phenomenon according to the Thomson model.

Such large deflections required strong forces to be acting, such as those between very small charged particles very close together. This would be possible, Rutherford reasoned, if all the positive charge, along with most of the atomic mass, were concentrated in a very small central region which Rutherford called the atomic *nucleus*. Now, if the α particle were also merely an atomic nucleus, the scattering problem could be treated by an analysis of the repulsion between two mass points that repel each other according to Coulomb's inverse square law. An α particle penetrating an atom near its edge (Fig. 21-8) would be deflected only a small amount; those passing closer to the nucleus would be repelled with a greater force and deflected through a larger angle. Rutherford, knowing the kinetic energy of his α particles, calculated that they would have to get to within about 10^{-12} cm from the center of the nucleus if they were to be turned back in the direction from which they came. Thus the radius of the nucleus could not be larger than this small distance. A mathematical analysis of the total deflections that an α particle might experience in penetrating a foil was in very close agreement with the experimental data gathered by Geiger and Marsden.

Since the radius of a whole atom is in the neighborhood of 10^{-8} cm, it is apparent that the volume of the atom must be largely made up of electrons arranged around the nucleus in such a way as to occupy this space. Because there would be a Coulomb force of attraction between the

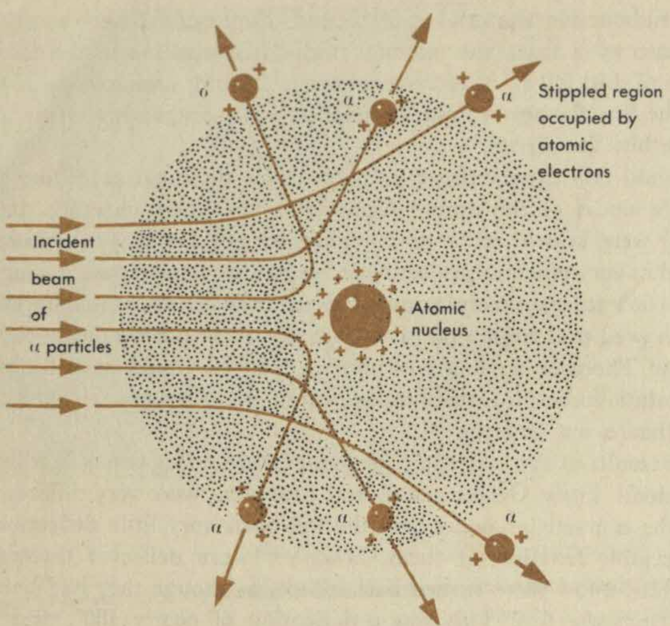


FIG. 21-8 The deflection of alpha particles by the repulsion of the small, massive, positively charged nucleus.

positive nucleus and the negative electrons, the two would be drawn together and the atom would vanish unless some provision were made to prevent it. It was suggested that the electrons might be orbiting rapidly around the nucleus, so that the electrostatic attraction would merely provide the necessary centripetal force. Thus in one bold stroke Rutherford transformed the static "watermelon model" of J. J. Thomson into a dynamic "planetary model" in which the nucleus plays the role of the sun and the electrons correspond to the individual planets of the solar system.

Because the mass of an electron is only about 1/7000 of the mass of an α particle, their contributions to the deflection of the α particles in Rutherford's scattering experiments could quite properly be ignored.

21-8 Conduction of Electricity in Solids

We have discussed the passage of an electric current through liquid solutions of acids, bases, and salts and through rarefied gases. In the first case, the current was due to the motion of positively and negatively charged ions, such as Ag^+ and NO_3^-, shouldering their way through the crowd of water molecules. In the second case, we dealt with positively charged ions flying in one direction and free negative charges, or electrons, flying in the opposite direction. But what happens when an electric current passes through solids, and why are some solids (all of them

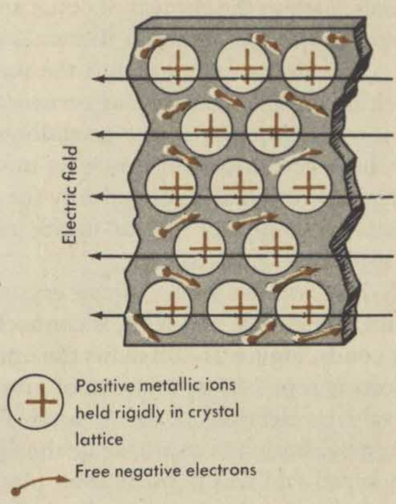


FIG. 21-9 The motion of free electrons explains the passage of electric current through metals.

classed as metals) rather good conductors of electricity while the rest of them, known as insulators, pass hardly any electric current at all? Since in solid materials all atoms and molecules are rigidly held in fixed positions and cannot move freely as they do in gaseous or liquid materials, the passage of electricity through solids cannot be due to the motion of charged atoms or atomic groups. Thus the only active electric carrier can be an electron, which, being much smaller than the atoms and molecules forming the crystalline lattice of a solid, should be able to pass between big atoms as easily as a small speedboat can pass through a heavily crowded anchorage of bulky merchantmen. Indeed, this is what takes place in metallic conductors. The high electrical conductivity of these substances is inseparably connected with the presence of free mobile electrons that rush to and fro through the rigid crystalline lattices (Fig. 21-9).

These electrons which belong to no particular atom play an essential part in holding the metallic crystal together. Indeed, without them the positive ions would all repel each other and the crystal would be immediately converted into a puff of vapor. Free as the electrons are to move about, wherever they go, they are surrounded by the positive ions and are attracted by all of them. They thus act as a sort of bonding device which gives the crystals the strength characteristic of most metals.

In nonmetals such as sulfur, each atom holds tightly to all its electrons, and the application of an electric field can cause nothing more than a slight deformation (electric polarization) of the atoms forming the crystal lattice. Such nonmetals must therefore in general be classified as nonconductors, or insulators.

21-9 Semiconductors

Some materials, such as the element silicon, cannot be classified as either good insulators or good conductors; thermal agitation of the atoms can knock loose a few electrons and permit the material to be slightly conductive. Such materials are known as *semiconductors*. A small amount of the proper kind of impurity in the crystalline structure of a semiconductor may, however, make it enormously more conductive. The three pictures in Fig. 21-10 explain how and why the presence of foreign atoms in the originally completely regular lattice may lead to such a large increase of electrical conductivity.

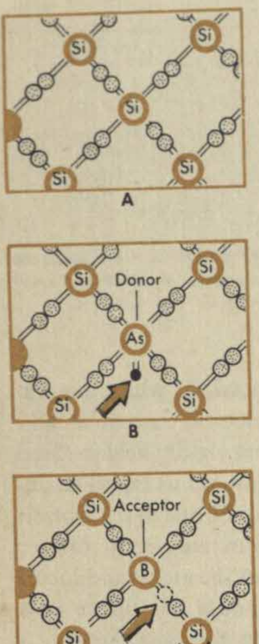


FIG. 21–10 A small proportion of the proper impurities in the crystal structure of a semiconductor makes it much more conductive.

In Fig. 21-10A, we see a pure silicon crystal in which each atom of silicon, having a chemical valence 4, is connected with four of its neighbors by four bonds. Figure 21-10B shows the situation that arises when one atom of silicon is replaced by an atom of arsenic, which has a valence of 5. Four valence electrons of the As atom form connections (bonds) with the four neighboring Si atoms, while the fifth "black sheep" electron is left unemployed and free to travel from place to place. The impurity atoms that give rise to free electrons in this way are known as *donors*. A reverse situation occurs when the Si atom is replaced by a trivalent atom of boron (Fig. 21-10C). In this case there will be a vacant place, or an *electron hole*, that breaks up the spotless regularity of the silicon crystal lattice. The impurity atoms that give rise to such "holes" are known as *acceptors*. A hole formed near a foreign atom present in the lattice may be filled up by an electron originally belonging to one of the neighboring silicon atoms, but in filling this hole the electron will leave a hole at the place where it was originally located. If this hole is filled by another neighboring electron, a new hole will move one step farther out (Fig. 21-11). Thus we can visualize the hole of that type as an "object" that is moving through the crystal, carrying a deficiency or negative charge, or, what is the same, a positive electric charge. Semiconductors that contain donor atoms and free electrons are known as *n-type* semiconductors, while those with acceptor atoms and holes are called *p-type* semiconductors (*n* and *p* stand for the negative or positive charge of the electric carriers). The electrical conductivity of *n*-type semiconductors is determined by the number of free electrons per unit valence and the ease with which they move through the crystal lattice, while in the case of *p*-type semiconductors it depends on the number and mobility of the holes.

21-10 Thermionic Emission

The principle of modern electronic tubes is based on the fact that red-hot metallic surfaces emit large numbers of free electrons. This phenomenon is in a way similar to the evaporation of liquids. Just as in the case of liquids, where the molecules are normally prevented from crossing the surface by mutual cohesive forces (surface tension forces), free electrons

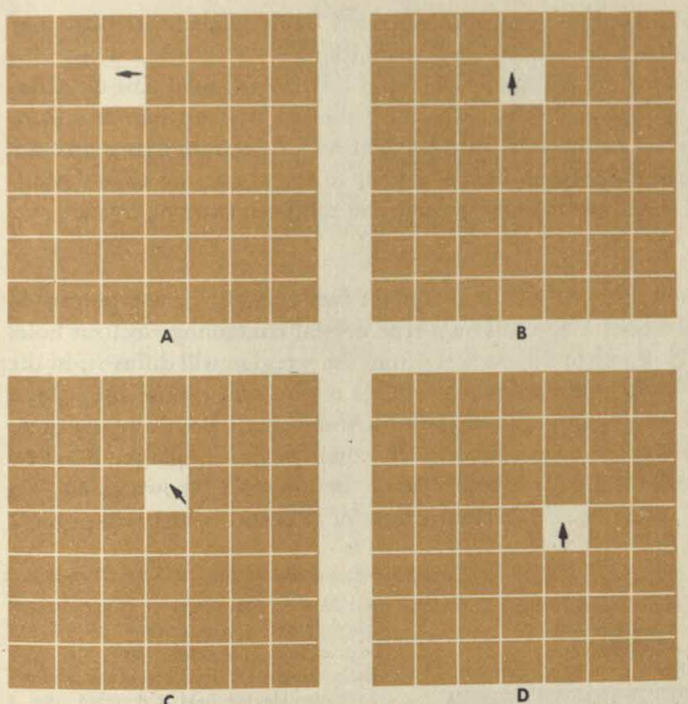


FIG. 21–11 The successive fillings of a "hole" by neighboring electrons (represented by shaded squares) makes the hole move (in this case downward and to the right as the filling electrons move upward and to the left).

are held inside the metal by electric attraction to the positive ions forming the lattice. But at sufficiently high temperatures, the kinetic energy of a small fraction of free electrons inside the metal becomes sufficiently high to overcome this surface barrier, and these electrons fly freely into space, to the great delight of physicists and radio engineers. The temperature at which the "evaporation" of electrons begins to be perceptible depends on the strength of the electric forces holding them in and is different for different metals. As in liquids, where the evaporation process goes easily in ether and alcohol, less easily in water, and quite slowly in heavy oils, we observe different rates of electron evaporation from different metals (fast for cesium, slower for tungsten, and quite slowly for platinum).

The operation of electron tubes in radio, TV, radar, etc. depends on a stream of electrons traveling to a positive anode from the hot cathode which emits them. This stream is controlled in different ways to provide the various functions of rectification (i.e., conversion of alternating into direct current), amplification, oscillation, and so on. They perform all these functions admirably well, but by far the greater part of the electrical energy used by such devices goes merely into heating the

cathodes so that they may emit electrons freely. This in its turn may create problems in cooling and ventilation.

This is one of the reasons for the rapid increase in "solid state electronics" applied to replace the electronic tubes in devices from transistor radios to giant computers. In these, instead of electrons flying through an evacuated space from the hot cathode to the anode, we have the migration of electrons or holes through the solid structure of crystals.

21-11 Crystal Rectifiers

Suppose now that we put into contact two crystals: an *n*-type crystal containing free electrons and a *p*-type crystal containing electron holes (Fig. 21-12). Some of the electrons from the *n* region will diffuse into the *p* region, while some holes from the *p* region will diffuse into the *n* region. Thus the *n*-type crystal will become slightly positively charged, and the *p*-type crystal will carry an equal negative charge. Between these opposite charges on both sides of the interface (known as an "*n-p* junction"), there will be an electric force of attraction which will prevent

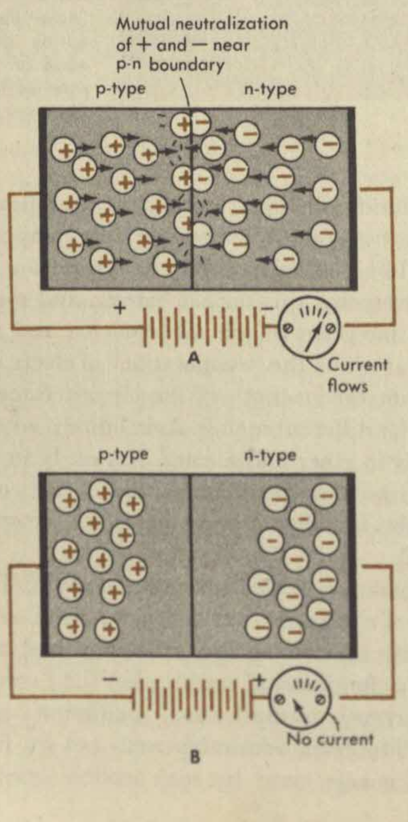


FIG. 21-12 The motion of electrons and holes across a *p–n* junction. (A) If the electric field is directed from *p* to *n*, a continuous current flows across the junction. (B) If the electric field is directed from *n* to *p*, no current can flow.

further diffusion, and the situation will be stabilized with a certain number of holes in the *n*-type crystal and an equal number of electrons in the *p*-type crystal. It must be remembered, however, that when free electrons and electron holes exist side by side in a given material, they can be mutually "annihilated" by a free electron filling a hole. In order to compensate for the losses due to this annihilation process, a small number of electrons and holes will continue to diffuse in opposite directions through the *n-p* junction.

Let us see what happens now if we apply an electric potential at the two ends of our crystal pair. If the positive pole of a battery is connected to the *p*-type crystal and the negative pole to the *n*-type crystal (Fig. 21-12A), there will be a force driving the holes to the right and the electrons to the left, and an electric current will begin to flow through the system. Since both crystals are now being invaded by holes and electrons crossing the border, the rate of mutual annihilation on both sides of the *n-p* junction will increase considerably, and more holes and electrons will have to be produced on both sides. These new electrons for the *n* type crystal will be supplied by electrons pouring through the wire from the negative pole of the battery, while new holes will be produced by electrons leaving the *p*-type crystal on their way to the positive pole of the battery.

If, on the other hand, we reverse the direction of the electric potential, the situation will be quite different (Fig. 21-12B). Now the electrons and the holes will be pulled in opposite directions, leaving a "no man's land" at the *n-p* junction. It is clear that under these conditions no current can flow through our double crystal. Thus we see that our device will conduct electric current in one direction but not in the opposite one. This property of one-way electric conductivity of *n-p* junctions permits us to use pairs of *n*-type and *p*-type crystals instead of hot-cathode vacuum tubes for the purpose of rectifying alternating currents.

21-12 Transistors

A thin layer of *p*-type crystals sandwiched between two *n*-type crystals (or, alternatively, an *n*-type between two *p*-types) can serve as an amplifier, and is in general called a *transistor*. Transistors cannot yet do all the things that can be done by some special types of vacuum tube, but their development will no doubt continue at a fast pace. The principal advantage of transistors over vacuum tubes lies in the fact that the controlled flow of electrons takes place entirely *within solid material*. Thus it is not necessary to use a large amount of power to keep a filament red-hot to "boil" electrons off into space. This, in addition to their simplicity, sturdiness, and small size, is rapidly causing transistors to take the place of vacuum tubes in many fields of electronics.

21-13 Solar and Radioactive Batteries

The properties of the *n-p* junction between two crystals can also be used for the direct transformation into electric energy of both solar radiation and the rays emitted by radioactive materials. The point is that when radiation is absorbed in the material of a semiconducting crystal, it knocks off some electrons from the atoms to which they belong, thus increasing the number of free electrons and electron holes. This increased number of electric carriers disturbs the electrostatic balance at the interface between the *n*- and *p*-type crystals and causes an electric current to run from the crystal containing acceptors to the crystal containing donors. One workable solar battery of this kind, developed in the laboratories of the Bell Telephone Company, consists of a silicon crystal with a slight arsenical contamination (donor) through its entire body, except for a thin upper *p*-type layer (one ten-thousandth of an inch thick), which is contaminated by boron and serves as an acceptor. The sun's rays that fall on the upper surface of this device are absorbed in the material of the crystal, produce extra electrons and extra electron holes, and stimulate an electron potential of about one-half volt. This device has about a 20 percent efficiency, and it produces a power of about 0.01 watt/cm^2 of its surface. A battery with a working surface of 10 m^2 (about 100 ft^2) installed on the roof of a house will produce a power of 100 watts, which, when stored in ordinary electric storage batteries, is sufficient to operate a 100-watt electric lightbulb for the same number of hours that the sun was shining during the day. Because of the present high cost of producing the elements of a solar battery, it would be highly irrational to use it for the purpose of saving on the electric bill; but such batteries will undoubtedly find many useful applications, one of which has already been the production of power for running the electric equipment in experimental satellites.

The principle of the solar battery can be used also for the direct transformation of α, β, and γ rays emitted by radioactive materials, such as fission products, into the energy of electric current. If such a device can be constructed with an efficiency comparable to that of the solar battery, the fission products that result from the operation of plutonium-producing piles and various nuclear power reactors could be used to run small household gadgets and devices employed in many other walks of life.

Questions

(21-1) **1.** In a cell, what is the name of the electrode with positive charge? with negative charge?

2. What are the ions, and the charge on each, into which (a) water decomposes, to a slight extent, (b) nitric acid decomposes in solution, (c) silver nitrate decomposes in solution? Give the charge on each ion.

(21-2) **3.** (a) Write the symbols, atomic numbers, and atomic weights of iron and potassium. (b) How does the mass of an iron atom compare with the mass of an atom of potassium? (c) What is the ratio of the mass of 7×10^{19} iron atoms to the mass of 7×10^{19} potassium atoms? (d) Consider a sample of 55.85 gm of iron, and one of 39.10 gm of potassium. How do the number of atoms in each of these samples compare?

4. Look up lead and zinc in Table 21-1, and (a) write their symbols, atomic number, and atomic weights. (b) What is the ratio of the mass of a lead atom to the mass of a zinc atom? (c) How does the mass of 3.42×10^{24} atoms of lead compare with the mass of an equal number of atoms of zinc? (d) How does the number of atoms in 20.719 gm of lead compare with the number of atoms in 6.537 gm of zinc?

5. Compare the number of atoms in 5 gm of iron to the number of atoms in 20 gm of zinc.

6. What is the number of atoms in 12 gm of lead and the number of atoms in 5 gm of zinc?

7. What is an equivalent weight of (a) hydrogen? (b) oxygen? (c) copper? (d) aluminum?

8. How many atomic weights and equivalent weights are there in (a) 50 gm oxygen? (b) 50 gm silver? (c) 50 gm aluminum?

9. An electrolytic cell contains a solution of gold chloride, and is placed in series with another cell containing a solution of a salt of copper (divalent). (a) By the time 0.60 gm of copper has been deposited, how much gold will have been deposited on the cathode of the other cell? (b) How many coulombs of charge will have flowed through the two cells?

10. An electrolytic cell contains a solution of a silver salt, and is connected in series with a cell containing a solution of a salt of zinc (divalent). (a) How many coulombs of charge must flow through the cells to deposit 10 gm of zinc? (b) In this time, how much silver will have been deposited?

11. Most of the copper used to make wire, etc., has been electrolytically refined by depositing it from a copper salt solution (divalent) onto a cathode. What is the cost of electrical energy required per kg of copper, if electricity costs the refiner 0.5 cent/KWH? (The cell operates at 0.2 volt.)

12. Molten salts (the pure salt, *not* a solution) are also ionized and can be electrolytically decomposed. Metallic sodium and chlorine gas can be produced by electrolysis of common salt (NaCl, both monovalent). How many pounds of sodium are produced for each pound of chlorine?

(21-4) **13.** What is the force on an electron in an electric field of 200 volts/cm?

14. An alpha particle is to be given a force of 4.8×10^{-15} nt in an electric field. What must be the strength of the field, in volts/cm?

15. What must be the flux density B, if a magnetic field is to exert a force of 3.2×10^{-10} dyne on an electron traveling 10^9 cm/sec^2 at right angles to the field?

16. What is the force on an electron whose speed is 10^{-9} cm/sec while moving at right angles to a field whose flux density is 3×10^{-3} weber/m^2?

17. What potential difference would be needed in the electron gun if an electron were to be undeviated in passing through an electric field of 300 volts/cm, and a perpendicular magnetic field of 10^{-3} weber/m^2?

18. If an alpha particle were to be accelerated by the same potential difference in the electron gun and passed through the same magnetic field as that of Question 17, what electric field (in volts/cm) would be needed if the alpha particle were to be undeviated? (An alpha particle is 7.30×10^3 times more massive than an electron.)

(21-5)

19. A droplet of oil has 3 extra electrons on it. What electric force is exerted on the droplet when it is in an electric field of 200 volts/cm?

20. Suppose, in a repetition of Millikan's oil-drop experiment, we have a drop 10^{-4} cm in diameter, with a density of 0.9 gm/cm^3. The plates are 2 cm apart, and a potential of 72 volts applied across them keeps the droplet just in balance. How many electronic charges are on the droplet?

(21-7)

21. The charge on a nucleus = the atomic number of the element $\times$ the charge of the electron; the charge on a neon nucleus, for example, is $10 \times 4.8 \times 10^{-10} = +4.8 \times 10^{-9}$ esu. (a) What is the potential energy of an alpha particle at the instant it has penetrated to 10^{-12} cm from a gold nucleus and is momentarily stationary? (b) What kinetic energy must be the particle have in order to penetrate to this distance from the nucleus?

22. An alpha particle has 1.2×10^{-6} erg of kinetic energy. What is the closest it can come to the nucleus of a silver atom? (See Question 21.)

23. Figure 21-8 is *not* drawn to scale. In the drawing, the diameter of the nucleus is about 0.3 inch. About what should the diameter of the atom be on this scale, if the drawing were to be in proper proportion?

chapter / twenty-two

The Energy Quantum

22-1 Light Emission by Hot Bodies

We know that in order for them to emit enough light to be seen, material bodies must be heated above a certain temperature. Hot radiators (100°C) emit radiation that we can feel with our cold hands in the winter, but none that our eyes can see. The heating elements of an electric range (at a low heat of about 750°C) glow with a faint reddish light that can be seen only if the kitchen is not too brightly illuminated. The filament of an electric lightbulb (about 2300°C) emits an intense yellow light. We can thus see qualitatively that the intensity of the radiation emitted by a heated body increases as its temperature increases; and that the wavelength of the most intense radiation shifts with increasing temperature from the red toward the blue end of the spectrum. Figure 22-1 shows how the intensity of radiation varies with wavelength at different temperatures.

Studies of experimentally determined curves such as these led to two fundamental laws governing radiation from hot solids and liquids.

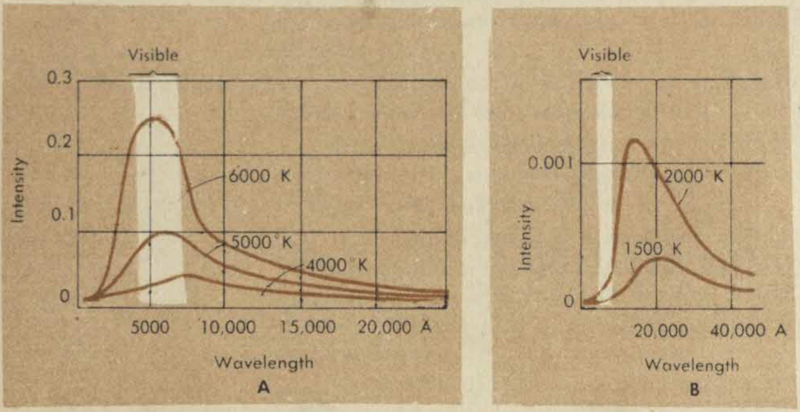


FIG. 22-1 Intensity distribution in the emission of radiation by bodies at different temperatures. The upper curve in (A) corresponds to the surface temperature of the sun—a large fraction of the total energy is in the visible part of the spectrum. The upper curve in (B)—notice that the scale is different from that of (A)—shows only a small fraction of visible radiation, most of which is red. In the 1500°K curve in (B) practically all the radiation is in the invisible infrared.

Wien's law: The wavelength of maximum intensity is inversely proportional to the absolute temperature of the emitting body.

In other words, the hotter the body, the shorter the wavelength. Wien's law can be put in quantitative numerical form by using the relationship $T\lambda_{max} = 0.29$ cm-deg. For the 100°C radiator (373°K) we find that $\lambda_{max} = 0.29/373 = 78 \times 10^{-5}$ cm or 78,000 Å. This wavelength at which the radiator radiates most intensely is 10 times as long as the longest deep-red wavelength of visible light; it should not be surprising that it does not glow in the dark!

The Stefan-Boltzmann law: The total rate (per unit area) of emission of energy of all wavelengths is directly proportional to the fourth power of the absolute temperature.

Quantitatively, this is $E = 5.67 \times 10^{-8}\ T^4$ watts/m². Our 100°C radiator can again serve as an example. For this relatively cool body, $5.67 \times 10^{-8} \times (3.73 \times 10^2)^4 = 1100$ watts/m² or 0.11 watts/cm². Note that this total rate of emission, calculated for different temperatures, is proportional to the areas under the different temperature curves in Fig. 22-1.

22-2 Infrared and Ultraviolet Radiation

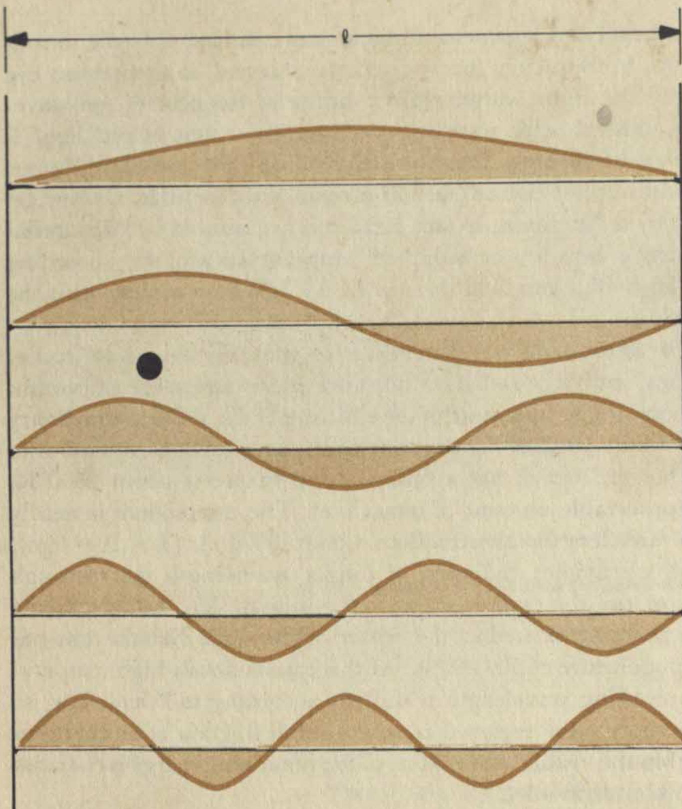
Just as in acoustical phenomena a human ear can hear only the sounds within a certain frequency (or wavelength) interval, so the human eye can see only the light within narrow limits of frequencies (or wavelengths). Radiation with wavelengths *longer* than that of red light is known as *infrared radiation*. It is also often called "heat radiation," since it is emitted from hot bodies (such as a room radiator) that are not yet hot enough to be luminous. In fact, heat rays are emitted by all material bodies no matter how low or high their temperatures are, but, according to the Stefan-Boltzmann law, their intensity falls very rapidly with the temperature.

Ultraviolet radiation has wavelengths *shorter* than the blue-violet end of the spectrum, and this radiation becomes more and more important with the increasing temperature of the emitting body. While the ordinary electric lightbulb (at 2300°C) does not emit any ultraviolet radiation to speak of, the sun, which has a surface temperature of about 6000°K, emits an appreciable amount of ultraviolet. The atmosphere is nearly opaque to wavelengths shorter than about 3000 Å (3×10^{-5} cm), but enough ultraviolet radiation of longer wavelength gets through to sunburn or tan the exposed parts of the human skin. As an extreme case, there is a star located in the center of the Crab Nebula that has a surface temperature of 500,000°K. At this tremendously high temperature, the prevailing wavelength is shifted, according to Wien's law, so far into the short-wave region that only a small fraction of its energy is emitted within the visible range. Most of the remaining energy is radiated in the invisible ultraviolet.

22-3 The Ultraviolet Catastrophe

During the last decade of the nineteenth century, a British physicist and astronomer, Sir James Jeans (1877–1946), made an attempt to treat the problem of the distribution of energy between different wavelengths of radiant energy in the same statistical way as was done by Maxwell in the case of the distribution of energy between different molecules of a gas. To do this, he considered radiant energy of different wavelengths enclosed in a cube, the walls of which are made of ideal mirrors reflecting a full 100 percent of any radiation falling on them. Of course, this so-called "Jeans's cube" is just an abstraction (since there are no such mirrors) and can be used only for the purpose of purely theoretical arguments; but in physics, we very often use idealized models of this sort.

In Fig. 22-2, we give a schematic picture of Jeans's cube and various waves that can exist within it. The situation is similar to that of the sound waves that can exist inside a cubical enclosure with perfectly reflecting walls, or to that of standing waves of any kind. The reflecting



FIG, 22–2 A view of "Jeans's cube," showing some of the standing waves that can exist (only horizontal waves are shown). The black dot is a coal-dust particle that can absorb waves of one wavelength and emit their energy at another wavelength.

walls must be nodes of the standing waves, so that l, the distance between the walls of the cube, is an integral number of half wavelengths.

The longest wave has a wavelength twice the length of the side l of the cube, and the next possible wavelengths are: $1, \frac{2}{3}, \frac{1}{2}, \frac{2}{5}$, etc. of l. We also assume that the box contains one or more "coal-dust particles" that are introduced here to permit the exchange of energy between the different modes of vibrations existing in the box. (These particles are supposed to absorb the energy from the vibrations of one wavelength and to reemit it at a different wavelength.)

We can now draw an analogy between the different vibrations within Jeans's cube and the molecules of gas contained in a similar cubical enclosure. Just as in the case of gas in which the total available kinetic energy can be distributed in various ways between individual molecules, the total available radiant energy within Jeans's cube can be distributed in various ways between the vibrations of different wavelengths. In an earlier chapter, we stated *the law of equipartition of energy*, according

to which all the molecules of gas share equally in the distribution of the total available energy, so that the mean energy of each molecule is simply equal to the total amount of energy divided by the number of molecules in the box. The same kind of statistical considerations led to the conclusion that the total radiant energy in the Jeans cube should be equally distributed between the vibrations of all different wavelengths. But here came a very serious difficulty! Whereas the number of molecules forming a gas, though very large, is still finite, *the number of possible vibrations in Jeans's box is infinite*, since we can continue the above given sequence of possible wavelengths beyond any limit. Thus, if the equipartition law holds in this case, as it certainly should, each individual vibration would get an infinitely small share of the total energy. Since, on the other hand, the sequence of wavelengths continues indefinitely in the direction of shorter and shorter wavelengths, *all the available energy will be concentrated in the region of infinitely short waves*. Thus, if we fill Jeans's cube with red light, it should rapidly become violet and then ultraviolet, turn into X rays and then into γ rays (such as are emitted by radioactive substances), and so on beyond any limit. What happens to radiant energy in the idealized case of Jeans's cube must also hold for the radiation in all practical cases, and the light emitted by red-hot pieces of coal in the fireplace should be turned into deadly γ rays even before it leaves the grate! Or, at least, that is what would happen if the laws of classical physics were applicable to radiant energy. This "Jeans's paradox," also known as the "ultraviolet catastrophe," gave a terrible blow to the self-satisfied classical physics of the nineteenth century and catapulted it into an entirely new field of thought and experience—now known as *quantum theory*—unprecedented in the history of physics. Although the advanced mathematics needed for detailed quantitative study of the quantum theory is not easy, the underlying concepts are not too difficult, even without extensive mathematics; and before finishing this chapter, the reader should acquire a general idea of what it is all about.

22-4 The Birth of the Energy Quantum

Just before the end of the last century, in Christmas week, 1899, at a meeting of the German Physical Society in Berlin, the German physicist Max Planck (1858–1947) presented his views on how to save the world from the perils of Jeans's ultraviolet catastrophe. His proposal was as paradoxical as Jeans's paradox itself, but it was much more helpful. According to Planck's view, the rays of the sun that pour into a room through the windows or the light of a table lamp do not represent a continuous flow of light waves, but rather a stream of individual *photons* (Fig. 22-3). A photon is a unit of electromagnetic radiation (you might imagine it as being a finite train of waves) having a certain

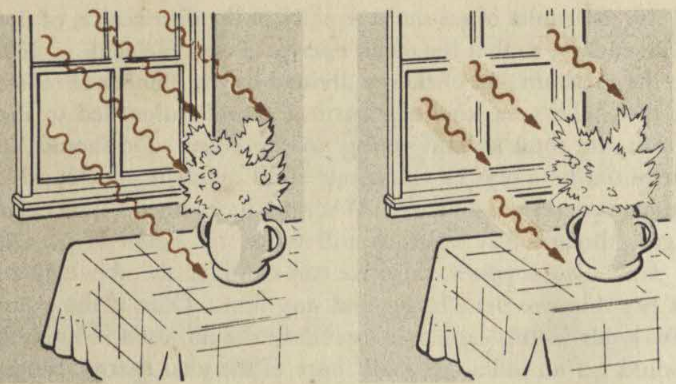


FIG. 22-3 (A) The old view of light as continuous wave trains whose amplitude increases with increasing intensity of the light; (B) the new view which shows light as a stream of "photons," each having a certain "quantum" of energy. The number of photons per second determines the intensity of the light.

wavelength and frequency, and having also a certain definite amount, or *quantum*, of energy.

To each photon frequency there corresponds a definite quantum of energy, and it is just as nonsensical to talk about three-quarters of a quantum of green light as it would be to talk about three-quarters of an atom of copper. Planck proposed that photons of different frequencies carry different quanta of energy, and that ***the energy of a photon is directly proportional to its frequency.*** Writing f for the frequency of the photon, and E for the quantum of energy it carries, we can express Planck's assumption as

$$E = hf$$

where h is a universal proportionality constant known as *Planck's constant*, and which has the value of 6.63×10^{-27} erg-sec or 6.63×10^{-34} joule-sec.

How does Planck's assumption of light quanta help to remove the troubles of the ultraviolet catastrophe? To understand this, let us look further into the consequences of the basic assumption that $E = hf$, that is, that radiant energy such as light flies about in packets of energy, the sizes of which are proportional to the frequency of the radiation. The long wavelength waves of radio have low frequencies and hence their quanta of energy are small. Visible light, with frequencies a billion times greater, comes in quanta whose energy is also a billion times greater. Energy must be absorbed and emitted in whole quanta, exactly —no fractional parts of quanta are allowed.

The difference between the energy demands of long-wave (low-frequency) and of short-wave (high-frequency) radiation has a very important effect on the application of the equipartition principle. Planck's

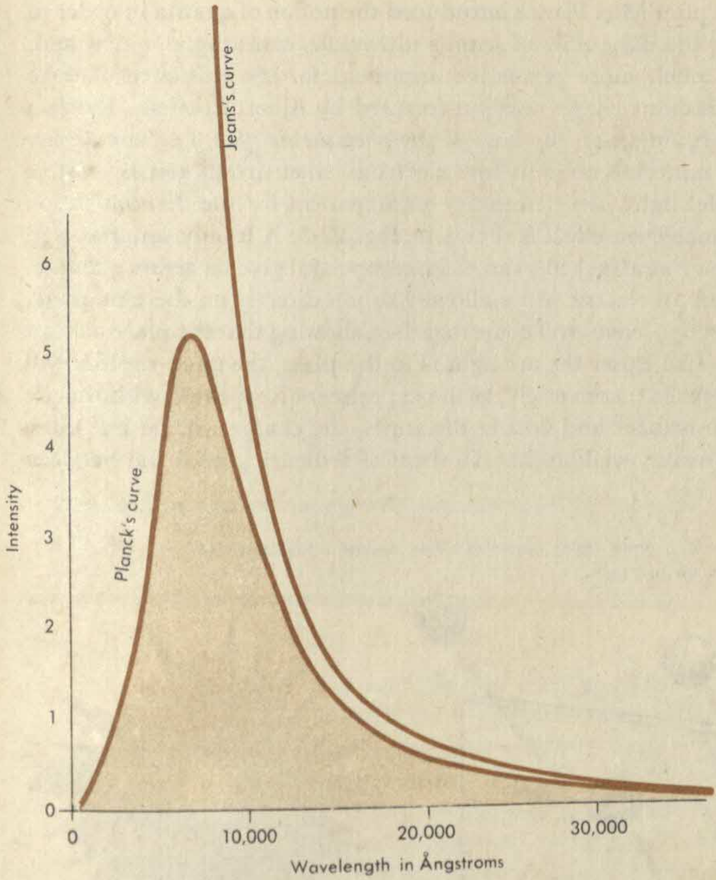


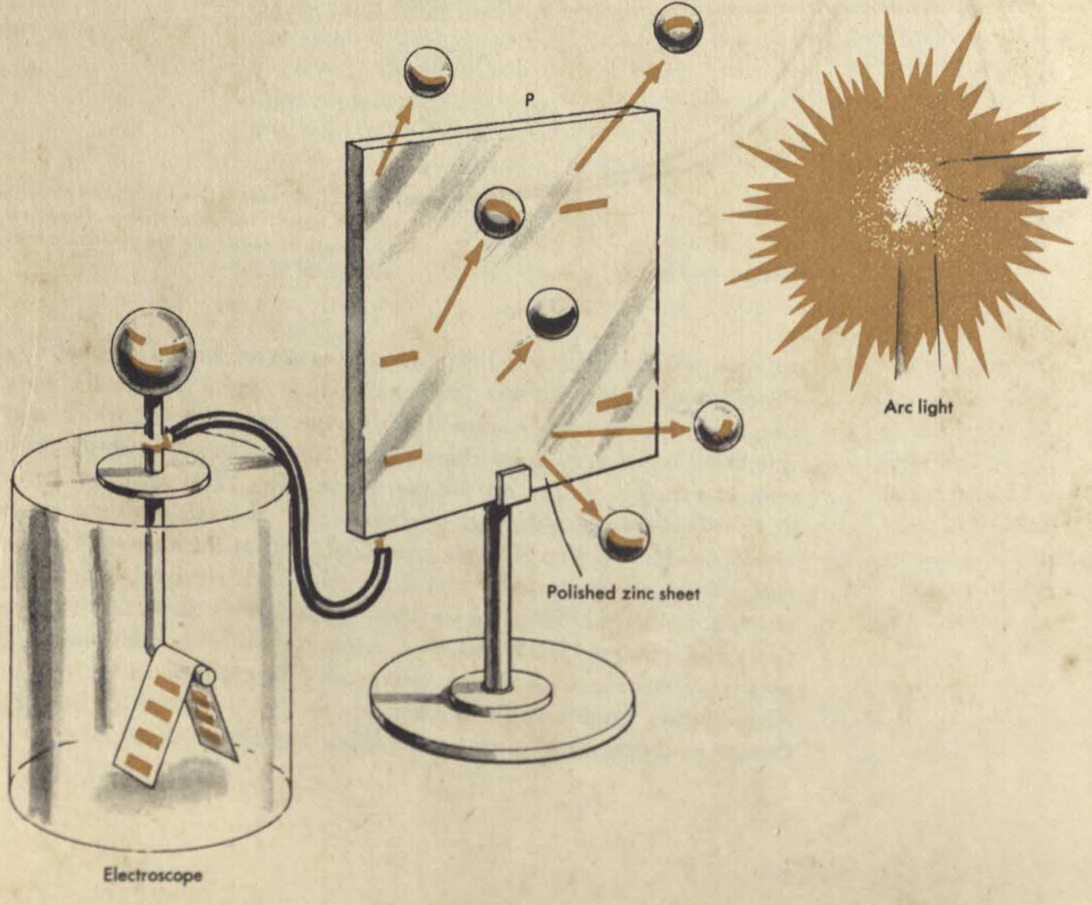
FIG. 22–4 Jeans's curve (wrong) and Planck's curve (right). These curves were theoretical attempts to derive the experimental curves shown in Fig. 22–1.

solution of the problem can be considered to take the form of a probability distribution: the vibrations with high demands have a very small chance of having their demands satisfied; low-frequency radiation, which asks but little, has a very good chance of getting it. In other words, both ends of Planck's energy distribution curve (Fig. 22-4) approach zero; at the short-wavelength, high-frequency end, the radiation has practically no chance of receiving a great deal; and at the long-wavelength end, the radiation stands a very good chance of receiving practically nothing. So instead of looking like the original Jeans's curve, the distribution curve obtained by Planck took a much more reasonable shape, in perfect agreement with the experimentally determined curves of Fig. 22-1. Further studies of the problem enabled Planck to derive from his theory the thermal radiation laws of Wien and Stefan-Boltzmann.

22-5 The Puzzle of the Photoelectric Effect

A few years after Max Planck introduced the notion of quanta in order to circumvent the difficulties of Jeans's ultraviolet catastrophe, a new and, in a way, much more persuasive argument for the existence of these packets of radiant energy was put forward by Albert Einstein. Einstein based his argument on the laws of the *photoelectric effect*, i.e., the ability of various materials to emit free electrons when irradiated by visible or ultraviolet light. An elementary arrangement for the demonstration of the photoelectric effect is shown in Fig. 22-5. A freshly sandpapered piece of zinc P is attached to an electroscope and given a *negative* charge. If light from an electric arc is allowed to fall directly on the zinc plate, the electroscope leaves will come together, showing that the plate has lost its charge. The closer the arc light is to the plate, the more rapidly will the charge be lost; conversely, as the experiment is repeated with the arc removed to greater and greater distances, the charge will be lost more slowly. However, we find that if a sheet of ordinary glass is put between

FIG. 22-5 A zinc plate losing electrons when exposed to ultraviolet radiation from an arc light.



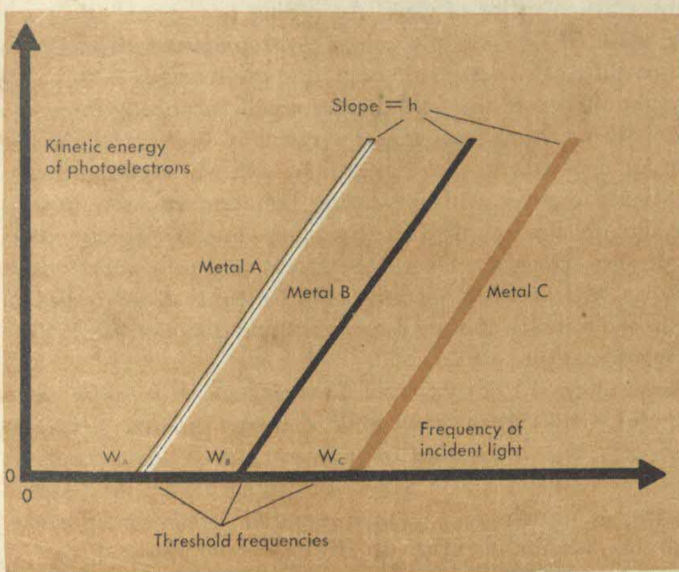


FIG. 22-6 The kinetic energy of emitted photoelectrons plotted against the frequency of the illuminating radiation, for three different metals.

the arc light and the zinc, the zinc will retain its negative charge, even if the arc is brought very close. Also, we find that if the zinc is originally given a *positive* charge, the arc light will have little apparent effect on the rate at which the charge is lost.

More careful experiments were performed, in which monochromatic radiation was used and in which the numbers and kinetic energies of the ejected electrons were carefully measured. These experiments gave results according to the following rules:

1. For a given frequency of incident radiation, the number of electrons produced by the photoelectric effect is directly proportional to the *intensity* of the radiation, while the velocity (or energy) of the emitted electrons remains the same, no matter what the intensity.

2. When radiation of different frequencies is used, the energy of the emitted electrons increases with increasing frequency. The graph of the kinetic energy of the emitted electrons against the radiation frequency is a straight line (Fig. 22-6).

3. When different metals are used, the graphs of Fig. 22-6 are different lines, but they all have the same slope. With radiation of a frequency less than a certain *threshold frequency* (different for different metals), *no* electrons will be emitted, no matter how great the intensity of the radiation.

These straightforward observations of the photoelectric effect presented great difficulties for the understanding of electromagnetic radia-

tion from the classical point of view. According to classical theory, a propagating wave of light or any other electromagnetic radiation is essentially a varying electromagnetic field, and the strength of the electric and magnetic forces in this field increases with increasing intensity. If the ejection of electrons from metal surfaces is due to the electric forces of the incident waves, then the energy with which the electrons are ejected should increase with increasing radiation intensity, just as seashore bathers are thrown off their feet more violently by big waves than by small ones. However, this quite natural conclusion does not fit the experimental facts, since an increase of the intensity of the radiation increases only the number of ejected electrons and does not affect their velocity or energy at all.

All this experimental evidence, said Einstein, could be quite satisfactorily explained by Planck's new idea of energy quanta. A certain amount of energy W_{Zn} is required to pull an electron loose from the attraction of the atoms in a zinc plate. According to the old classical theories, the energy of light or ultraviolet radiation spread out in spherical waves, so the amount of energy an electron could absorb from one tiny spot on such a spreading wave front would be negligible. However, Einstein argued, the old classical picture does not represent what actually happens. The entire energy of the photon is absorbed in one bite by a single electron. Planck's relationship $E = hf$ tells how much energy will be in a quantum of any given frequency f.

For the zinc plate, W_{Zn} is greater than the energy associated with a photon of visible light, so that no matter how much visible light shines on the plate, no electron will receive enough energy to break loose. This is the situation with a sheet of glass screening the arc light—ordinary window glass shuts out the invisible but highly energetic ultraviolet radiation. With the glass removed, the ultraviolet radiation from the arc, being of higher frequency than visible light and hence of proportionally higher energy, is absorbed by the electrons in the plate. The energy of a photon of ultraviolet is greater than W_{Zn}, so the electrons can escape, carrying any left-over excess energy with them in the form of kinetic energy. With the plate negatively charged, the departing electrons are repelled, and the plate gradually loses its charge. A positively charged plate, however, will attract the electrons back as quickly as they escape, so there is in this case no loss of charge.

For a general energy relationship, we need consider three terms: hf, which is the entire energy of the quantum absorbed by the electron; W, the energy required to pull the electron free from the surface; and $\frac{1}{2}mv^2$, the kinetic energy of the electron as it leaves. Simple consideration of the conservation of energy gives us

$$hf - W = \tfrac{1}{2}mv^2.$$

Now, if we graph f against $\frac{1}{2}mv^2$, as in Fig. 22-6, we should expect a

straight line. (Remember that h and W are both constants.) Planck's constant h determines the slope of the line, which will therefore naturally be the same for all metals. The bottom ends of the lines on the drawing represent electrons with zero kinetic energy—i.e., electrons just barely pulled away from the attractions of their neighbors.

For an electron to be pulled off without anything left over for kinetic energy, the hf of the absorbed quantum must equal W, and the frequency at which this occurs is the *threshold frequency*. The threshold frequencies and threshold wavelengths for some representative metals are as follows:

Platinum	$\lambda = 1980$ Å,	or	$f = 1.51 \times 10^{15}$ hz
Silver	$\lambda = 2640$ Å,	or	$f = 1.13 \times 10^{15}$ hz
Potassium	$\lambda = 7100$ Å,	or	$f = 4.22 \times 10^{14}$ hz

In his classical paper on this subject, published in 1905, Einstein indicated that the observed laws of the photoelectric effect can be understood if, following the original proposal of Max Planck, one assumes that ***electromagnetic radiation propagates through space in the form of individual energy packages called photons, and that, on encountering an electron, such a photon communicates to the electron its entire energy.***

This revolutionary assumption explains quite naturally the observed fact that an increase of the intensity of light leads to an increase of the number of photoelectrons, but not of their energy. More intense light means that more light quanta of the same kind will fall on the surface per second, and since a single photon can eject one and only one electron, the number of electrons must increase correspondingly. On the other hand, by decreasing the wavelength of incident light we increase the frequency and, consequently, the amount of energy carried by each individual photon, so that in each collision with a free electron in the metal these photons will communicate to it a correspondingly larger amount of kinetic energy.

As an example, let us investigate a little further the behavior of the zinc plate in Fig. 22-5. From a handbook, we can find that for zinc, W (called the photoelectric *work function*) is about 6.4×10^{-12} erg. This is the energy it must absorb if an electron is to be just barely pulled away, with nothing left over for kinetic energy; and this energy must be absorbed from a single photon. From $E = hf$, it can be determined that for a photon to have this much energy it must have the frequency

$$f = \frac{E}{h} = \frac{6.4 \times 10^{-12}}{6.63 \times 10^{-27}} = 0.96 \times 10^{15} \text{ hz}.$$

For the corresponding wavelength, we go to $v = f\lambda$ and find its wavelength to be $\lambda = 3 \times 10^{10}/0.96 \times 10^{15} = 3.1 \times 10^{-5}$ cm, or 3100 Å.

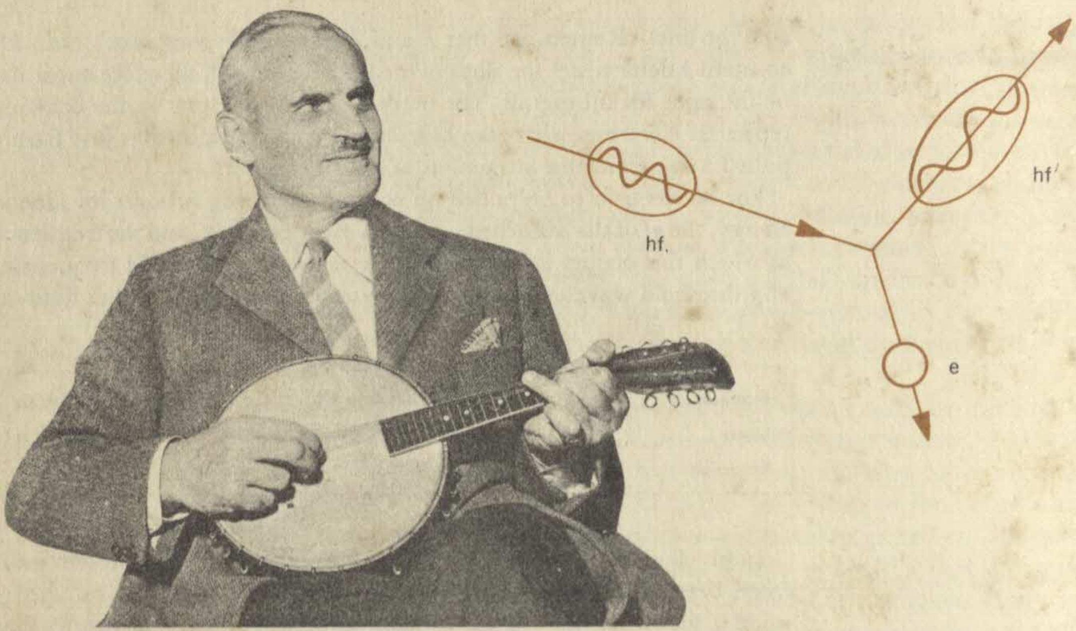
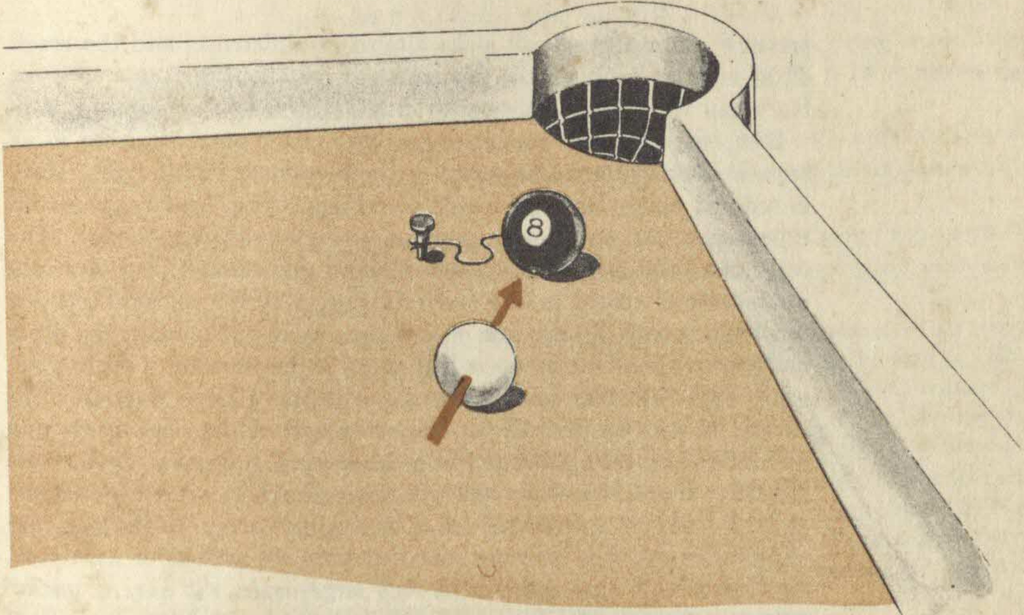


FIG. 22-7 Dr. Arthur H. Compton, who visualized a collision between a photon and an electron as analogous to a collision between elastic balls. The photon shown has less energy after its collision, and a correspondingly longer wavelength.

This is well beyond the visible into the ultraviolet. Radiation of wavelength longer than this will have energy quanta too small to pull electrons free, and none will be emitted, no matter how high the intensity; shorter wavelengths will have more than enough energy, and the excess will result in kinetic energy of the electrons emitted.

22-6 The Compton Effect

The Planck-Einstein picture of individual energy packages, or light quanta, forming a beam of light and colliding with the electrons within matter intrigued the mind of an American physicist, Arthur Compton (Fig. 22-7) who, being of a very realistic disposition, liked to visualize collisions between photons and electrons as similar to those between the ivory balls on a billiard table. He argued that, in spite of the fact that the electrons forming the planetary system of an atom are bound to the central nucleus by attractive electric forces, these electrons would behave almost as if they were completely free if the photons which hit them carried sufficiently large amounts of energy. Suppose that a black ball (electron) is resting on a billiard table and is bound by a thread to a nail driven into the table's surface and that a player, who does not see



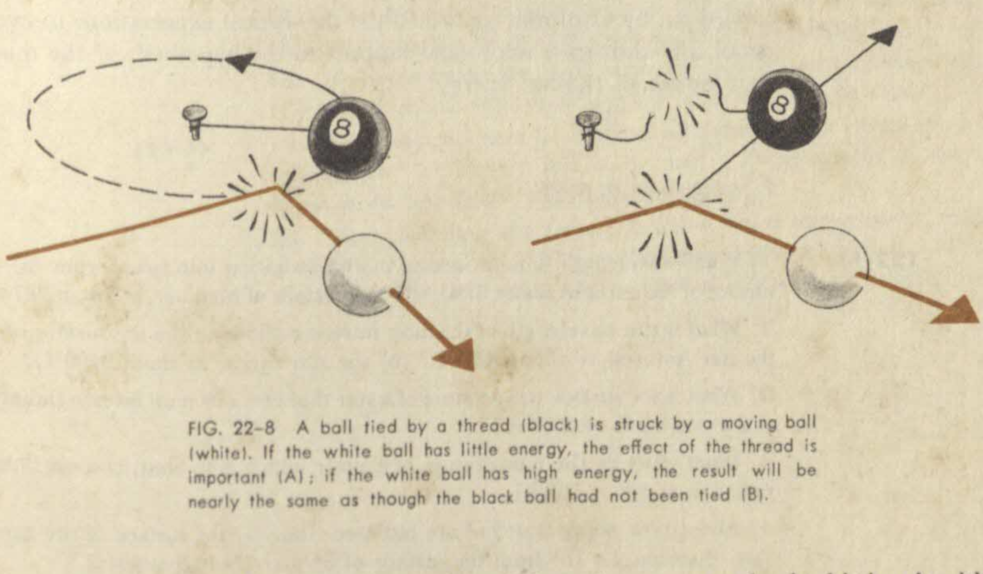


FIG. 22-8 A ball tied by a thread (black) is struck by a moving ball (white). If the white ball has little energy, the effect of the thread is important (A); if the white ball has high energy, the result will be nearly the same as though the black ball had not been tied (B).

the thread, is trying to put it into the corner pocket by hitting it with a white ball (photon) (Fig. 22-8). If the player sends his ball with a comparatively small velocity, the thread will hold during the impact and nothing will come from this attempt. If the white ball moves somewhat faster, the thread may break, but in doing so it will cause enough disturbance to send the black ball in a completely wrong direction. If, however, the kinetic energy of the white ball exceeds, by a large factor, the work necessary to break the thread that holds the black ball, the

presence of the thread will make almost no difference, and the result of the collision between the two balls will be about the same as if the black ball were completely unbound.

Compton knew that the binding energy of the outer electrons in an atom is comparable to the energy of the photons of visible light. Thus, in order to make the impact overpoweringly strong, he selected for his experiments the energy-rich photons of high-frequency X rays. The result of a collision between X-ray photons and (almost) free electrons can be indeed treated in very much the same way as a collision between two billiard balls. In the case of an almost head-on collision, the black ball (electron) will be thrown at high speed in the direction of the impact, while the white ball (X-ray photon) will lose a large fraction of its energy. In the case of a side hit, the white ball will lose less energy and will suffer a smaller deflection from its original trajectory. In the case of a mere touch, the white ball will proceed almost without deflection and will lose only a small fraction of its original energy. In the language of light quanta, this behavior means that in the process of scattering, ***the photons of X rays deflected through large angles will have a smaller amount of energy and, consequently, a longer wavelength.*** The experiments carried out by Compton confirmed the theoretical expectations in every detail, and thus gave additional support to the hypothesis of the quantum nature of radiant energy.

Questions

(22-1)

1. What wavelength is most intense in the radiation into space from (a) the surface of the earth at about 27°C? (b) the surface of Mercury, at about 327°C?

2. What is the wavelength of the most intense radiation from the surface of (a) the star Antares, at about 2600°C? (b) the star Sirius, at about 9700°C?

3. What is the surface temperature of a star that emits its most intense radiation at a wavelength of 2000 Å?

4. What must be the temperature of a plate which is to emit its most intense radiation at 1.2×10^4 Å?

5. About how many watts/m^2 are radiated from (a) the surface of the earth? (See Question 1.) (b) from the surface of Sirius? (See Question 2.)

6. How does the rate of energy emission per square mile from the surface of Sirius compare with the rate from the surface of Antares? (See Question 2.)

7. The voltage applied to a lamp filament is adjusted until its temperature is 1500°K, at which point the filament draws 25 watts. The voltage is now increased until the filament draws 50 watts. What is the filament temperature now?

8. The temperature of an electric heater element is 1200°K when it draws 200 watts. What must its wattage be in order to raise its temperature to 1800°K?

(22-4) **9.** What is the energy of a quantum of (a) a radio wave, frequency 870 kilocycles/sec? (b) green light, $\lambda = 5500$ Å? (c) an X ray, $\lambda = 0.6$ Å?

10. What is the energy of a quantum of (a) a radio wave whose frequency is 1490 kilocycles/sec; (b) an infrared photon of $\lambda = 2 \times 10^4$ Å; (c) a gamma-ray photon, $\lambda = 6 \times 10^{-3}$ Å?

11. If a photon is to have an energy of 10^{-12} erg, what must be its wavelength?

12. What is the wavelength of a photon whose quantum of energy is 3×10^{-15} erg?

(22-5) **13.** How much energy is needed to pull an electron away from a sheet of silver?

14. How much energy is needed to pull an electron away from an atom of platinum?

15. Yellow light ($\lambda = 5890$ Å) falls on a potassium surface. (This must be done in a vacuum to avoid oxidation of the potassium by the air.) (a) How much work is needed to pull an outer electron free from a potassium atom? (b) How much energy is carried by a photon of this yellow light? (c) With what kinetic energy are electrons ejected from the potassium surface? (d) With what speed?

16. A sheet of silver is illuminated by monochromatic ultraviolet radiation of $\lambda = 1810$ Å. (a) How much work is needed to free an electron? (b) How much energy is carried by each incident photon? (c) How much kinetic energy can a photoelectrically emitted electron have in this experiment? (d) With what maximum speed are electrons emitted?

17. What frequency of radiation must fall on a silver surface in order that the photoelectrons may be emitted with a speed of 10^8 cm/sec?

18. If photoelectrons are to be emitted from a potassium surface with a speed of 6×10^7 cm/sec, what frequency of radiation must be used?

(22-6) **19.** An X-ray photon of $\lambda = 1.00$ Å is scattered by an electron in a Compton interaction. The electron flies away with a kinetic energy of 4×10^{-10} erg. What is the wavelength of the scattered photon?

20. The wavelength of a photon is 1.50 Å before scattering by an electron in a Compton interaction. The scattered photon's wavelength is 1.54 Å. What is the kinetic energy of the scattered electron? What is its velocity?

chapter / twenty-three

The Bohr Atom

23-1 Bohr's Orbits

When Rutherford (at that time just plain Ernest Rutherford and not yet Sir Ernest or Lord Rutherford) was at the University of Manchester performing his epoch-making experiments that demonstrated the existence of the atomic nucleus, a young Danish physicist named Niels Bohr (1885–1962) came to work with him on the theoretical aspects of the atomic structure problem. Bohr was highly impressed by Rutherford's new atomic model, in which the electrons revolved around the central nucleus in very much the same way as the planets revolve around the sun, but he could not understand how such a motion could be at all possible in an atom. Whereas the planets of the solar system are electrically neutral, electrons are charged with negative electricity. As was mentioned in the discussion of oscillating circuits, energy in the form of electromagnetic radiation is radiated from accelerated electric charges. It is apparent from Maxwell's development of electromagnetic theory (which is too mathematical for us to discuss in detail here) that this is a fundamental part of the way electric charges behave. And an

electron revolving rapidly about a nucleus must have a constant centripetal acceleration. According to the well-established classical theory, this acceleration must cause a constant loss of energy by radiation, so that the electrons in the Rutherford model would be certain to spiral in toward the central nucleus and fall into it when all their rotational energy was spent on radiation.

Bohr calculated that the emission of electromagnetic waves would cause the electrons forming an atomic system to lose all their energy and fall into the nucleus within one hundred-millionth of a second! Thus, on the basis of conventional mechanics and electrodynamics, the planetary system of electrons revolving around the atomic nucleus as visualized by Rutherford could not exist for more than an extremely short period of time. This was in direct contradiction of the fact that atoms *do exist permanently* and do not show any tendency to collapse. How could it possibly be? Bohr's solution of this conflict between the conclusions of conventional mechanics and the facts of nature was straightforward and just: Since nature cannot be wrong, conventional mechanics must be wrong, *at least when applied to the motion of electrons within an atom.* In making this revolutionary statement concerning the motion of electrons within an atom, Bohr followed the precedent established by Planck and Einstein, who had some time before declared that the good old Huygens light waves were not what they were supposed to be according to the conventional views, but rather a bunch of individual oscillating photons.

It is always much easier to say that something is wrong than to find a way to make it right, and Bohr's criticism of conventional mechanics in the case of atomic electrons would be of no value whatsoever if he could not show a way out of the difficulty. The way he proposed was so odd and unconventional that he kept the manuscript locked in his desk for almost two years before deciding to send it in for publication. When this revolutionary paper finally appeared in the year 1913, it sent out a shock wave of amazement through the world of contemporary physics.

Defying the well-established laws of classical mechanics and electrodynamics, Bohr stated that in the case of the motion of electrons within an atom, the following postulatory rules must strictly hold:

I. ***From all the mechanically possible circular and elliptical orbits of electrons moving around the atomic nucleus, only a few highly restricted orbits are "permitted," and the selection of these "permitted" orbits is to be carried out according to specially established rules.***
II. ***Circling along these orbits around the nucleus, the electrons are "prohibited" from emitting any electromagnetic waves, even though conventional electrodynamics says they should.***

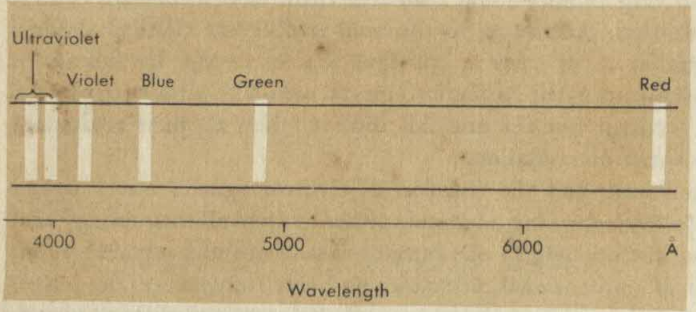


FIG. 23-1 Part of the spectrum of hydrogen atoms, in or near the visible range, indicating the color of the emitted frequencies.

III. Electrons may "jump" from one orbit to another, in which case the energy difference between the two states of motion is emitted in the form of a single Planck-Einsteinian quantum.

The whole thing sounded quite incredible, but it *did* permit Bohr to interpret the regularities of spectra emitted by various atoms and to construct a consistent theory of internal atomic structure.

It had been known for many years that the chemical elements in the form of gas or vapor emitted radiation of only certain specific frequencies when excited by either heat or a strong electric discharge. When looked at through a spectroscope, a tube of glowing gas shows a spectrum that consists only of bright colored lines against a black background. Each bright line is the image of the spectroscope slit, formed by radiation of one definite wavelength. Each element has its own pattern of colored lines, and the spectrum of hydrogen is especially simple—in the visible range it has four lines, as shown in Fig. 23-1. In 1885, a German school teacher, J. J. Balmer, discovered that the frequencies of these lines were given by a very simple formula. This formula states that the frequencies of the observed lines are exactly proportional to the difference between the inverse square of 2, i.e., $\frac{1}{4}$, and the inverse squares of 3, 4, 5, etc. Calculating these differences, we obtain

$$\frac{1}{2^2} - \frac{1}{3^2} = \frac{1}{4} - \frac{1}{9} = 0.138889$$

$$\frac{1}{2^2} - \frac{1}{4^2} = \frac{1}{4} - \frac{1}{16} = 0.187500$$

$$\frac{1}{2^2} - \frac{1}{5^2} = \frac{1}{4} - \frac{1}{25} = 0.210000$$

$$\frac{1}{2^2} - \frac{1}{6^2} = \frac{1}{4} - \frac{1}{36} = 0.222222$$

etc.

Multiplying these values by 3.289×10^{15}, we get*

4.569×10^{14}
6.168×10^{14}
6.908×10^{14}
7.310×10^{14}
etc.

which is to be compared with the values

4.569×10^{14}
6.168×10^{14}
6.908×10^{14}
7.310×10^{14}

representing the actually measured frequencies of hydrogen's spectral lines. (The frequency itself cannot be directly measured. The wavelength is determined by measuring the angle of deviation in a grating spectrometer.)

In 1906, a Harvard spectroscopist, Lyman, discovered that the spectrum of hydrogen also included a series of lines in the ultraviolet, and that their frequencies could be given by a similar formula:

$$f(\text{Lyman}) = 3.289 \times 10^{15} \times \left(\frac{1}{1^2} - \frac{1}{n^2}\right)$$

in which n took on successive integral values of 2, 3, 4, . . .

Two years later, a German spectroscopist, Paschen, found another hydrogen spectrum series in the infrared, with frequencies

$$f(\text{Paschen}) = 3.289 \times 10^{15} \times \left(\frac{1}{3^2} - \frac{1}{n^2}\right)$$

in which n took integral values of 4, 5, 6, . . .

Experimental observations such as the above were something that any successful theory of the atom would have to explain, and this was the job Bohr tackled.

Bohr's original restrictions concerning the motion of the electron in a hydrogen atom pertained strictly to the case of circular motion and required that the angular momentum of the electron be an integral multiple of $h/2\pi$, where h is Planck's constant. For a small particle like an electron in an orbit, the expression for angular momentum becomes quite simple. The angular momentum for any body is $I\omega$, and for a single particle the moment of inertia I is mr^2. The angular speed of rotation ω we can express as v/r, so the angular momentum of the electron is simply mrv. Bohr's theory, then, requires that

* The reason for this particular factor is explained by Bohr's work later in this chapter.

$$mrv = \frac{nh}{2\pi}$$

where n can be any integer.

The centripetal force keeping the moon circling the earth is provided by the gravitational attraction between moon and earth; for the electron in a hydrogen atom, the centripetal force is the electric attraction between it and the positively charged proton that constitutes the nucleus. Both proton and electron have a charge that we can call e esu; and at a separation of r cm, the electric attraction will be e^2/r^2. If we set this equal to centripetal force, we get the expression

$$\frac{e^2}{r^2} = \frac{mv^2}{r}$$

or

$$e^2 = mv^2r = mrv \times v.$$

We can now substitute for mrv the value required above from Bohr's "quantization" of angular momentum and get

$$e^2 = \frac{nhv}{2\pi}$$

from which

$$v = \frac{2\pi e^2}{nh}.$$

From the expression for angular momentum, we see that

$$r = \frac{nh}{2\pi mv}.$$

Replacing v by its value as determined above, we get

$$r = \frac{nh \times nh}{2\pi m \times 2\pi e^2}$$

$$= n^2 \times \frac{h^2}{4\pi^2 me^2}.$$

We know that Planck's constant h equals 6.63×10^{-27} erg-sec; the mass of an electron, m, is 9.11×10^{-28} gm; and e, the charge on a proton or electron, is 4.80×10^{-10} esu. From these known values, we can compute the permitted radii of the hydrogen atom. The smallest radius r_1 will be determined when $n = 1$:

$$r_1 = \frac{(6.63 \times 10^{-27})^2}{4\pi^2 \times 9.11 \times 10^{-28} \times (4.80 \times 10^{-10})^2}$$

$$= 5.3 \times 10^{-9} \text{ cm}$$

$$= 0.53\text{Å}.$$

Larger orbits will be possible when n (known as the "quantum number")

is given the successively larger integral values of 2, 3, 4, etc., to describe other possible orbits:

$$r_2 = 4r_1, \quad r_3 = 9r_1, \quad r_4 \doteq 16r_1, \quad \text{etc.}$$

For the normal smallest value r_1, the diameter of a hydrogen atom is about 1 Å. This compares reasonably well with the 4-Å diameter we previously estimated for the larger air molecule, using an entirely different method of approximation.

Let us consider the energies of the electron in various of its permitted orbits. In any orbit, the energy will be in two parts, kinetic and potential. The kinetic energy can be readily figured from the value of v we derived above:

$$\text{KE} = \frac{1}{2}mv^2 = \frac{1}{2}m\left(\frac{2\pi e^2}{nh}\right)^2$$
$$= \frac{1}{n^2} \times \frac{2\pi^2 me^4}{h^2}.$$

We can arbitrarily take the potential energy of an electron to be zero when it is separated from its proton by an infinitely large distance, so that, as the electron is brought closer to the proton, its potential energy will be negative. This assumption need not worry anyone; the potential energy of a large weight on a shelf in the laboratory will also be negative if we happen to have chosen the roof of the building as our zero level. We are concerned only with *energy changes*, in any case, and these will be the same no matter what reference we select to be zero. The advantage of this particular selection of the zero energy level is that it makes the algebra easy, and it allows us to express the total energy in a very simple form.

It was shown in an earlier chapter that the electric potential at a point r cm distant from a charge Q esu is

$$V = \frac{Q}{r} \text{ ergs/esu.}$$

The potential energy of a charge q esu placed at this point will be Qq/r ergs. From this, we find that the potential energy of an electron r cm from a proton will be

$$\text{PE} = -\frac{e^2}{r}$$
$$= \frac{1}{n^2} \times \frac{-4\pi^2 me^4}{h^2}.$$

The total energy of the electron, PE + KE, will thus be, for a quantum number n,

$$E = \frac{1}{n^2} \times \frac{-4\pi^2 me^4}{h^2} + \frac{1}{n^2} \times \frac{2\pi^2 me^4}{h^2}$$
$$= \frac{1}{n^2} \times \frac{\pi^2 me^4}{h^2}(-4+2)$$
$$= -\frac{1}{n^2} \times \frac{2\pi^2 me^4}{h^2}$$
$$= -\frac{1}{n^2} \times \frac{2\pi^2 \times 9.11 \times 10^{-28} \times (4.80 \times 10^{-10})^4}{(6.63 \times 10^{-27})^2}$$
$$= 2.18 \times 10^{-11} \times \frac{1}{n^2} \text{ erg.}$$

23-2 Radiation and Energy Levels

Bohr's third postulate states that when an electron jumps from one level to another level of lower energy, the *energy difference* ΔE is radiated as a photon of energy hf: $\Delta E = hf$. Let us consider as an example an electron that jumps from the $n = 3$ level to the $n = 1$ level in a hydrogen atom. Then

$$hf = \Delta E = E_3 - E_1$$
$$= -2.18 \times 10^{-11} \times \left(\frac{1}{3^2} - \frac{1}{1^2}\right)$$

and

$$f = \frac{2.18 \times 10^{-11}}{6.63 \times 10^{-27}} \times \left(\frac{1}{1^2} - \frac{1}{3^2}\right)$$
$$= 3.289 \times 10^{15} \times \left(1 - \frac{1}{9}\right).$$

Thus we see that Bohr's work gave a convincing theoretical explanation for the puzzling multiplier 3.289×10^{15} which had to be put in as a purely empirical factor by Balmer, Lyman, and Paschen. And, to continue,

$$f = 3.289 \times 10^{15} \times 0.888$$
$$= 2.921 \times 10^{15} \text{ hz}$$

and

$$\lambda = \frac{c}{f} = \frac{3 \times 10^{10}}{2.921 \times 10^{15}}$$
$$= 1.027 \times 10^{-5} \text{ cm} = 1027\text{Å}.$$

We see that this line, well into the ultraviolet, is a member of the Lyman series.

It is easy to go from this specific example to a general formula covering all the lines of the hydrogen emission spectrum:

$$f = 3.289 \times 10^{15} \times \left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right)$$

in which n_2 can be any integer, and n_1 can be any integer greater than n_2. For $n_2 = 1$, we get the Lyman series; for $n_2 = 2$, we get the Balmer series; for $n_2 = 3$, we get the Paschen series; for $n_2 = 4, 5, 6, \ldots$, we get other series of lines of generally lower frequency and longer wavelength in the infrared (Fig. 23-2).

A question naturally arises about how the electron in the hydrogen atom could get up to a higher energy level in order to jump back and emit energy. Obviously, the electron can get in a higher energy orbit only by absorbing energy. This absorbed energy may come from collisions if the gas is heated to a high temperature. It may come from the energy of an electric spark or cathode-ray tube discharge, or *it may arise from the gas absorbing, from radiation falling on it, those same frequencies that it is able to emit.*

Since an atomic electron can occupy only those particular energy levels (or orbits) that we have just calculated, the energy differences between levels are the same whether the electron emits radiation or absorbs it. It cannot absorb a quantum which would bring it to an energy between permitted levels, and it cannot absorb a part of a quantum from a passing photon; it must be the whole quantum or nothing. Thus the frequencies of the absorbed photons must be exactly the same as the frequencies of the photons that the atom can radiate.

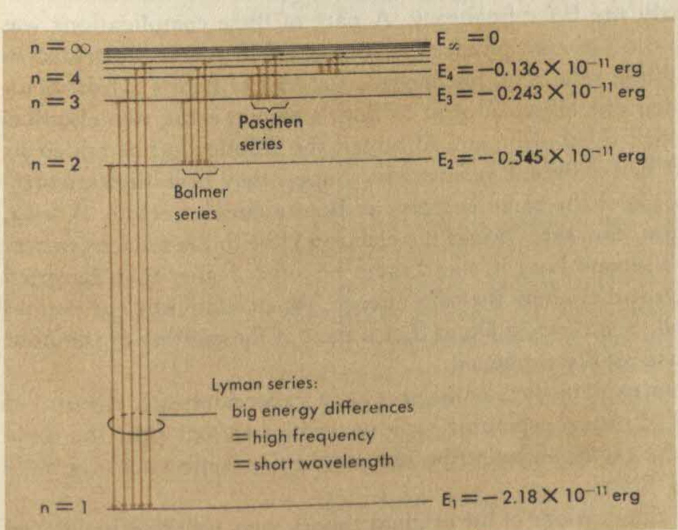


FIG. 23-2 A schematic diagram of Bohr's hydrogen atom model, showing the origin of different series of lines in the hydrogen spectrum.

23-3 Successes and Limitations of the Bohr Theory

So far as we have seen, the Bohr theory of the hydrogen atom seems to be a very successful one. But does this highly artificial picture of light emission by a hydrogen atom really make any sense? Haven't Bohr's postulates been specially adjusted so as to lead in the end to the empirically established Balmer's formula? Certainly they were! But this is exactly how a new theory is usually introduced in physics. Newton introduced the notion of universal gravity in order to interpret the observed motion of the moon around the earth and of the planets around the sun, and in the very same way Bohr introduced his three postulates pertaining to electron motion in an atom, and light emission by "jump" processes, in order to interpret the observed laws of hydrogen line spectra. However, the criterion for the validity of any new theory in physics is not only that this theory should give a correct interpretation of previous observations but that it also *predict* things which can be later confirmed by direct experiment.

To a limited extent, Bohr's original theory was able to do this. Other spectral series were discovered in the infrared, just as predicted. The theory was also successfully applied to other one-electron systems with great accuracy. (Ionized helium is such a system. Helium has two electrons, but if one of these is ionized off, the resulting He^+ is very like H, except that its nucleus has a charge of $2e$. Doubly ionized lithium, Li^{++}, is another.)

With the development of highly precise spectroscopy, however, complications began to present themselves. A Balmer series line, for example, is not a single line, but *six* lines, separate and distinct, but all of very nearly the Bohr frequency. A part of these complications was resolved by the German physicist A. Sommerfeld and is illustrated in Fig. 23-3. While retaining, unchanged, the first of Bohr's orbits, Sommerfeld added one elliptical orbit to Bohr's second orbit, two elliptical orbits to Bohr's third orbit, etc. Although the elliptical orbits added by Sommerfeld had different geometrical shapes, they nevertheless corresponded to *almost* the same energies as Bohr's circular orbits. A long, narrow ellipse, however, brings the electron close to the nucleus, where (by Kepler's second law) its speed must be much higher than its speed in a circular orbit of about the same energy. When relativistic corrections were applied, Sommerfeld found that a part of the splitting of the Bohr lines was successfully explained.

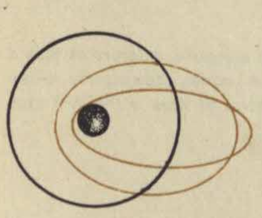


FIG. 23-3 Circular and elliptical atomic orbits of nearly identical energy, according to Sommerfeld.

For further explanation, it was necessary to consider each electron as a little ball of charge, spinning rapidly on its own axis (like the daily rotation of the earth) and contributing an additional effect not originally included by Bohr.

All these elaborations of the original theory were well and good, and

would still permit the atom to be pictured as a system of definite ball-like charged particles revolving in orbits about the massive central particle of the nucleus. This became a difficult picture to imagine, however, when it was found that some "orbits" must have *zero* angular momentum. The only orbit that could have zero angular momentum would have been an oscillation back and forth along a straight line penetrating to the center of the nucleus. This seemed a most unlikely situation.

And although the Bohr model worked well for the main features of the hydrogen atom, it would not give correct answers for atoms with more than one electron. The theory could not be adjusted or patched up to make it yield the frequencies actually observed in even the two-electron atom of helium.

As we shall see, this picture had to be abandoned. It is still, however, a great milestone in scientific achievement, and one that is still very useful. It showed that Planck's quantum ideas were a necessary part of the atom and its inner mechanism; it introduced the idea of quantized energy levels and explained the emission or absorption of radiation as being due to the transition of an electron from one level to another. As a model for even multielectron atoms, the Bohr picture is still useful. It leads to a good, simple, rational ordering of the electrons in larger atoms and qualitatively helps to predict a great deal about chemical behavior and spectral details. Even the most abstract theorist, although he now ignores the Bohr model in his calculations, still *imagines* the atom as composed of electrons in various orbits and finds this picture to be a very useful thing.

Questions

(23-1)

1. What is the frequency of the second line in the hydrogen Lyman series? of the fourth line?

2. Calculate the frequencies of the third and fifth lines in the Paschen series of hydrogen.

3. Since both the electron and the proton in a hydrogen atom have mass, there is also a gravitational attraction between them. What is the ratio of their gravitational to their electrical attraction? Do you think this justifies the gravitational force having been neglected in computing the Bohr orbits? (The mass of a proton is about 1840 × the electron mass.)

4. Show that the dimensions of Planck's constant (erg-sec) are equivalent to the dimensions of angular momentum in the CGS system.

5. What is the radius of the third quantum orbit of the electron in a hydrogen atom?

6. What is the radius of the fourth permissible orbit in a hydrogen atom?

7. Consider a singly ionized helium ion. (It has only one orbital electron, and its nucleus has a charge of $+ 2e$.) What is the radius of the smallest possible orbit for this electron? (The derivation will be the same as for hydrogen, except that at the beginning, instead of $e \times e = e^2$, you will have $2e^2$. With this knowledge, you can jump immediately to the end and make the necessary change in the expression for hydrogen.)

8. The nucleus of the lithium atom has a charge of $+3e$, and the atom has three orbital electrons. For a doubly ionized lithium atom (i.e., it has only one orbital electron left), what is the radius of the smallest electron orbit?

9. What is the total energy of the electron in its second permitted orbit in an ionized helium atom? (See Question 7.)

10. What is the total energy of the electron in its third permitted orbit in a doubly ionized lithium ion? (See Question 8.)

11. What is (a) the kinetic energy, (b) the potential energy, of the $n = 1$ orbit of ionized helium? (See Question 7.)

12. What is (a) the kinetic energy, (b) the potential energy, of the $n = 1$ orbit of doubly ionized lithium? (See Question 8).

(23-2)

13. (a) What is the wavelength of the photon emitted when a hydrogen electron jumps from $n = 2$ to $n = 1$? (b) Is this a photon of visible light? (c) What is the name of the series of which this spectral line is a member?

14. A hydrogen atom makes a transition from $n = 3$ to $n = 2$. (a) What is the wavelength of the emitted photon? (b) Is this photon visible? (c) Of which spectral series is this photon a member?

15. A singly ionized helium ion (see Question 7) has a spectrum qualitatively similar to the spectrum of atomic hydrogen. Transition from higher n to $n = 1$ will produce a series analogous to the Lyman series of hydrogen. How will their frequencies compare with those of the hydrogen Lyman series?

16. Doubly ionized lithium ions (see Question 8) produce a spectrum similar to that of hydrogen atoms. Transitions from higher n to $n = 2$ will give a spectral series analogous to the Balmer series of hydrogen. How will the frequencies of these lines compare to the Balmer frequencies?

17. Consider a sample of hydrogen in which most of the atomic electrons are in their $n = 1$ state. White light passed through this gas will show absorption lines where some frequencies have been absorbed. Which series will show the stronger absorption, the Balmer or the Lyman?

18. The lower dense layers of a star's atmosphere emit a continuous spectrum of all frequencies; in passing through the rare upper layers of the star's atmosphere, frequencies are absorbed characteristic of the elements in the upper atmosphere, which always contains much hydrogen. In the spectrum of cool stars (up to 5000°K) the Balmer absorption lines are very weak. They become darker in stars of higher temperature up to about 10,000°K. In hotter stars, the Balmer lines become weaker again. Explain.

chapter / twenty-four

The Structure of Atoms

24-1 Quantum Numbers

Having become acquainted with the possible orbital motions of the one hydrogen electron, as permitted by the Bohr-Sommerfeld rules, we can now tackle the question of the structure of the multielectron atoms.

Bohr introduced the *principal quantum number n*. In his model, this was done by quantizing, or assigning only certain specific values to, the angular momentum of the electron. However, this did not last long. We still keep *n* as the principal quantum number, but it now refers to the general energy level of the electron, which is also quantized, or allowed only certain values, as we have previously calculated.

The second quantum number *l*, the *orbital quantum number*, came into the picture as a result of Sommerfeld's introduction of elliptical orbits. This quantum number, rather than *n*, is the measure of the angular momentum of the electron, which depends on the ellipticity of the orbit. A circular orbit has the maximum angular momentum; an oscillation along a straight line penetrating the nucleus has a zero angular momen-

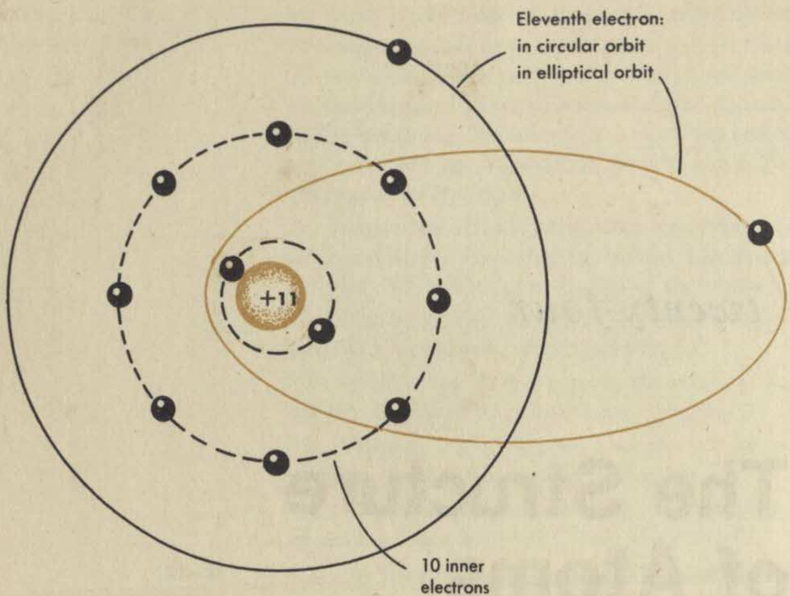


FIG. 24–1 Schematic diagram of a sodium atom: the energy of the outermost eleventh electron depends on whether its orbit is circular or elliptical.

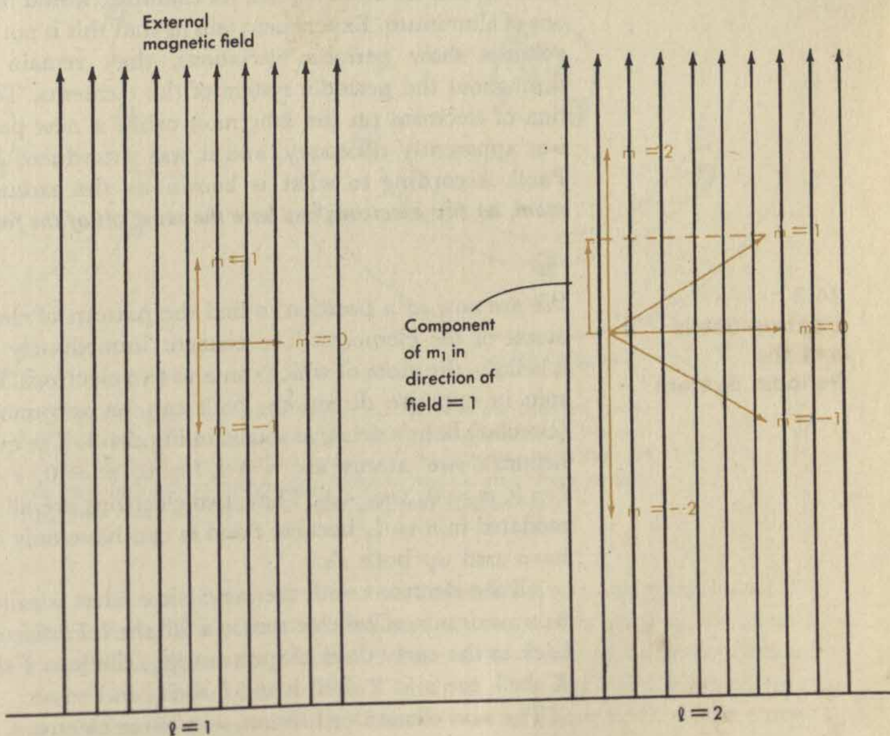
tum. It was found (through the analysis of spectra of various elements) that l could have only integral values ranging from 0 to $n-1$. For $n=1$, for example, l can have only the value 0; for $n=3$, l can be 0,1, or 2. For the hydrogen atom, the shape of the orbit has little effect; except for the small relativistic correction, the energies are all the same for a given value of n. In larger atoms, though, the effect may be considerable. Figure 24-1 is a schematic diagram of a sodium atom, which has 11 electrons and whose nucleus therefore has a charge of $+11e$. The eleventh electron, when in a circular orbit, is constantly shielded from the nucleus by the negative charges of the 10 inner electrons. In an elliptical orbit, it penetrates this shield and spends part of each orbital revolution close to the strong electrostatic pull of the nucleus. The energies of electrons when in orbits of different l may therefore be quite different, even though n is the same.

Spectroscopists had observed that many normally single spectral lines split up into several lines when the gas emitting them was in a strong magnetic field, and a way was found to interpret this, too, in terms of the Bohr model. We may look on an orbiting electron as being a tiny electric current which produces a little magnetic field perpendicular to the plane of the orbit. It was found that the electronic magnetic field (and hence

also the orientation of the orbit) could not be pointed in just any direction. This, too, is quantized: the magnetic field of the electron, measured in units we shall not worry about, must be pointed so that its component in the direction of an external magnetic field has only integral values. Since the electron's magnetic field (technically, its *magnetic moment*) is proportional to l, these permitted positions may be determined schematically as shown in Fig. 24-2. For $l = 1$, m can have the integral values 1, 0, -1; for $l = 2$, m can have the integral values 2, 1, 0, -1, -2; and so on for higher values of l. In general, there are $(2l + 1)$ possible values for m. (When $l = 0$, m too can of course have only the value $m = 0$.)

In connection with the fine splitting of spectral lines, we have already mentioned the idea of electron spin. Still holding to the picture given by the Bohr model, we can imagine the electron as a rotating ball of negative charge, which would also create a small magnetic moment for the electron itself, regardless of its orbit. In order to explain spectral

FIG. 24–2 Vectors representing the orientation of an electron's magnetic moment in a magnetic field. Only those directions are allowed that give an integral component (+, −, or 0) in the direction of the field.



details, it was necessary to conclude that the spin, too, must be quantized. The vector representing its magnetic moment, parallel to its axis of rotation, must be aligned either in the direction of an external magnetic field, or directly against it. This gives rise to the fourth (and last!) of the quantum numbers: s, which can be either $+\frac{1}{2}$ or $-\frac{1}{2}$. (The "$\frac{1}{2}$" comes into the scheme because the magnetic moment of the electron spin is just half of that arising from each unit of l.)

We have seen that in the case of the hydrogen atom, the electron puts itself in the lowest energy level available; if it is excited into a higher energy, it at once falls back to $n = 1$, either all in one jump or in several steps, emitting its excess energy as radiation as it does so. Is this also the case for the multielectron atoms?

If it were, the normal, unexcited state for any atom would find all its electrons (as many as 103 in the case of lawrencium!) packed together in a very overcrowded ring along the first of Bohr's orbits. With atoms of increasing atomic number, this ring would become more and more crowded because it would have to accommodate more electrons and also because its radius would become smaller and smaller owing to the stronger electric attraction exercised by the central nucleus.

If this were true, the size of atoms would decrease rapidly with atomic number, and an atom of lead, for example, would be much smaller than one of aluminum. Experiment tells us that this is not so; although atomic volumes show periodic variations, they remain roughly the same throughout the periodic system of the elements. To avoid this congestion of electrons on the innermost orbit, a new postulatory restriction was apparently necessary, and it was introduced by the physicist W. Pauli. According to what is known as the *exclusion principle*, ***in any atom, no two electrons can have the same set of the four quantum numbers.***

24-2 Electron Shells and the Periodic System

We are now in a position to find the pattern of electron motion in the atoms of the elements. The element immediately following hydrogen is helium, the atom of which contains two electrons. If these two electrons spin in opposite directions, both can be accommodated on the first (circular) Bohr's orbit, as shown in Fig. 24-3. The quantum numbers for helium's two atoms are $n = 1, l = 0, m = 0, s = -\frac{1}{2}$; and $n = 1, l = 0, m = 0, s = +\frac{1}{2}$. These two electrons are all that can be accommodated in $n = 1$, because l and m can have only the value 0, and we have used up both s's.

All the electrons with the same n are what is called a *shell*; there can be a maximum of $2n^2$ electrons in a full shell. For historical reasons dating back to the early days of spectroscopy, the $n = 1$ shell is known as the K shell, the $n = 2$ shell is the L shell, and so on.

The next element is lithium, with three electrons. (It is the number of

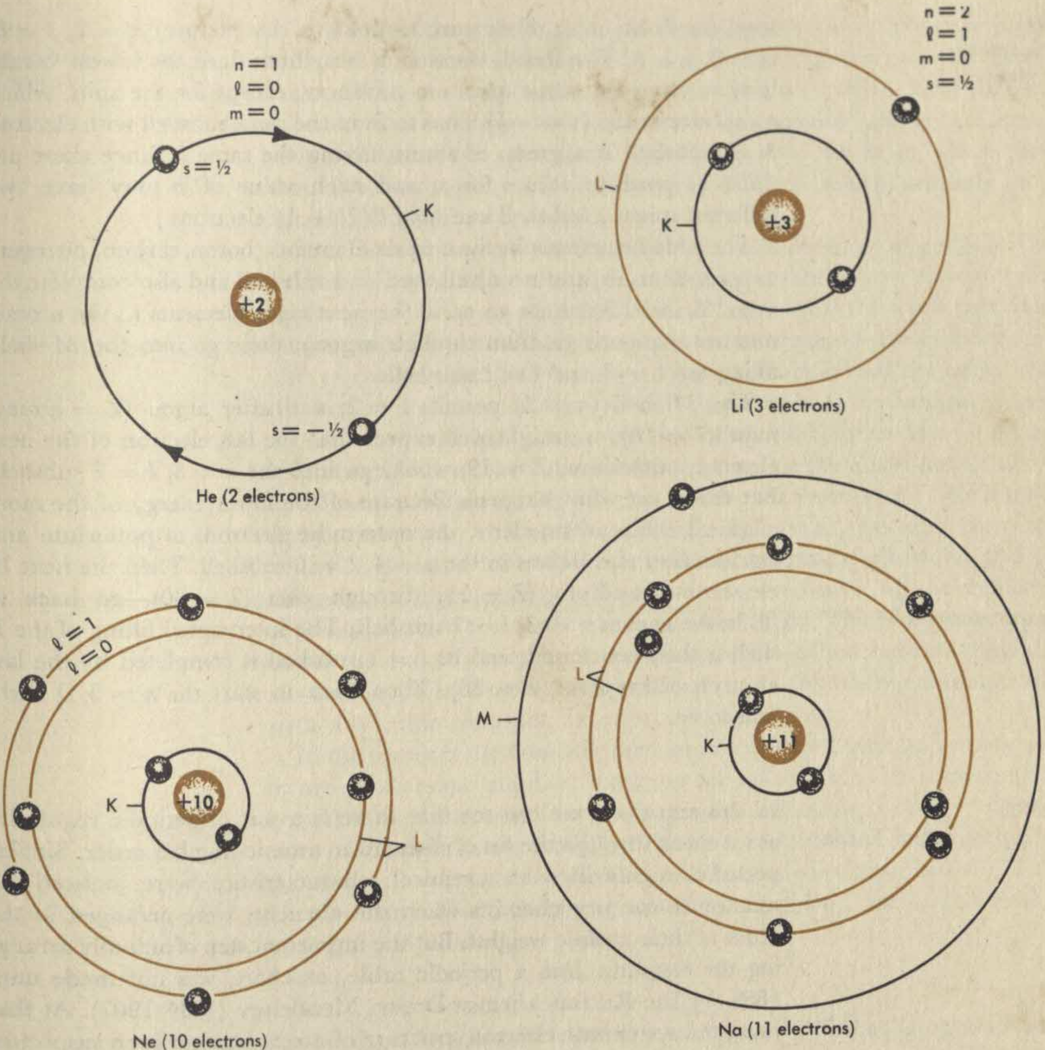


FIG. 24-3 How the electron shells and subshells are filled in atoms of increasing atomic number. As the atomic number increases, the size of the shells becomes smaller because of the increased attraction of the greater nuclear charge. The result is that all atoms have roughly the same size, no matter how many electrons they contain.

electrons which gives each element its *atomic number*, as given for all the elements in Table 21-1. Since atoms are electrically neutral, the atomic number also gives the positive charge of the element's nucleus.) Because there is no more room in the K shell ($n = 1$), the third electron must be content with a place in the L shell ($n = 2$). There it will fit the energy

level (or Bohr orbit, if we care to hold to this picture) $n = 2$, $l = 0$, $m = 0$, $s = \frac{1}{2}$. The fourth electron of beryllium finds the lowest vacant level to have the same quantum numbers, except for the spin, which must be reversed ($s = -\frac{1}{2}$) if it is to fit in the same subshell with electron 3. (A subshell is a group of atoms having the same l. Since there are $(2l + 1)$ possible values for m and each value of m may have two different spins, a subshell can hold $2(2l + 1)$ electrons.)

The added electrons for the next six elements (boron, carbon, nitrogen, oxygen, fluorine, and neon) fill the $l = 1$ subshell and also complete the $n = 2$ L shell. So as we go on to the next eight elements in the atomic number sequence (sodium through argon), these go into the M shell, filling the $l = 0$ and $l = 1$ subshells.

The M shell ($n = 3$) permits $l = 2$, and after argon (Z = atomic number = 18), we might well expect that the last electron of the next element, potassium, Z = 19, would go into the $n = 3$, $l = 2$ subshell. But this is *not* what happens. Because of the lower energy of the more elliptical orbits of smaller l, the outermost electrons of potassium and calcium find themselves in the $n = 4$, $l = 0$ subshell. Then the next 10 elements—scandium (Z = 21) through zinc (Z = 30)—go back to fill in the vacant $n = 3$, $l = 2$ subshell. The interrupted filling of the N shell is then continued, and its $l = 1$ subshell is completed by the last electron of krypton (Z = 36). Then back to start the $n = 5$ O shell, and so on.

24-3 The Periodic Table

In this sequence, we can see that there is a sort of periodic regularity as we check through the list of elements in atomic-number order. Similar periodic regularities in chemical characteristics were noticed by nineteenth-century chemists when the elements were arranged in the order of their atomic weights. But the important step of actually arranging the elements into a periodic table, or chart, was not made until 1869, by the Russian chemist Dmitri Mendeleev (1834–1907). At that time, the systematic electron structure of atoms was not even suspected, and Mendeleev was further handicapped by the fact that the list of chemical elements then known was rather incomplete.

Driven by a strong belief that there *must be* a regular periodicity in the natural sequence of elements, Mendeleev made the bold hypothesis that the deviations from the expected periodicity in his list were due to the failure of contemporary chemistry to have discovered some of the elements existing in nature. Thus, in constructing his table, he left a number of empty spaces to be filled in later by future discoveries, and in certain instances he reversed the atomic-weight order of the elements in order to comply with the demands of the regular periodicity of their chemical properties.

Mendeleev thought his needed reversals of atomic-weight order were evidence that some of the atomic weights had been erroneously determined. This was not the case, however, and we can see now that his difficulty was caused by the fact that the fundamental property of atoms which dictates their chemical behavior and the order in which they should be arranged is *not* their atomic weights (which are only incidental), but their atomic numbers.

Such a table in modern, complete form is shown in Fig. 24-4. The horizontal rows, or *periods*, include elements whose outer electrons are in the same shell—or what is the same thing, have the same principal quantum number n. The vertical columns of elements are more important to the chemist. The first two columns on the left of the table, and the last six columns on the right, include the major chemical *groups*, all of whose members have very similar chemical behavior. In group I, for example, we have in addition to hydrogen the alkali metals, all of low density, chemically very active, and with a valence of 1. We should not be surprised to find them chemically similar, because chemical behavior is controlled very largely by the outermost electrons, and in this group all the elements have just one outer electron. All the members of group II have two outer electrons, and so on. The last group, from helium to radon, comprise the noble gases, so-called because they are aloof and separate from the others, forming no stable chemical compound with any other elements, or even with themselves.

In the center of the table are the *transition elements*. Here, as we progress in order of atomic number, electrons are being added in buried shells beneath the surface of the atom. In period 6, for example, as we progress from left to right in the transition elements, electrons are being added to the $n = 5$, $l = 2$ subshell. In this same period 6, elements $Z = 58$ to 71 (the lanthanides, shown below the main table), are added in the $n = 4$, $l = 3$ subshell.

24-4 Chemical Valence and Bonds

An appreciation of the electronic structure of different elements leads us to a simple explanation of their chemical valence. Atoms which are just beginning a new shell find it very easy to get rid of their outer electrons. And as can be seen from the structure of the noble gases, 8 outer electrons (representing filled $l = 0$ and $l = 1$ subshells) is a very stable configuration; atoms having only a few less than 8 outer electrons find it energetically favorable to add electrons to complete this quota of 8. For example, chlorine ($Z = 17$) has 2 electrons in the first shell, 8 in the second, and 7 in the third, which makes the outer shell short 1 electron. On the other hand, a sodium atom ($Z = 11$) has 2 electrons in the first shell, 8 in the second, and only 1 electron as the beginning of the third shell. Under these circumstances, when a chlorine atom encounters

PERIODIC CHART OF THE ELEMENTS

METALS

Transition Elements

Metalloids and Non-metals

PERIODS	I	II											III	IV	V	VI	VII	VIII
1	1 H Hydrogen																	2 He Helium
2	3 Li Lithium	4 Be Beryllium											5 B Boron	6 C Carbon	7 N Nitrogen	8 O Oxygen	9 F Fluorine	10 Ne Neon
3	11 Na Sodium	12 Mg Magnesium											13 Al Aluminum	14 Si Silicon	15 P Phosphorus	16 S Sulfur	17 Cl Chlorine	18 Ar Argon
4	19 K Potassium	20 Ca Calcium	21 Sc Scandium	22 Ti Titanium	23 V Vanadium	24 Cr Chromium	25 Mn Manganese	26 Fe Iron	27 Co Cobalt	28 Ni Nickel	29 Cu Copper	30 Zn Zinc	31 Ga Gallium	32 Ge Germanium	33 As Arsenic	34 Se Selenium	35 Br Bromine	36 Kr Krypton
5	37 Rb Rubidium	38 Sr Strontium	39 Y Yttrium	40 Zr Zirconium	41 Nb Niobium	42 Mo Molybdenum	43 Tc Technetium	44 Ru Ruthenium	45 Rh Rhodium	46 Pd Palladium	47 Ag Silver	48 Cd Cadmium	49 In Indium	50 Sn Tin	51 Sb Antimony	52 Te Tellurium	53 I Iodine	54 Xe Xenon
6	55 Cs Cesium	56 Ba Barium	57 La Lanthanum	72 Hf Hafnium	73 Ta Tantalum	74 W Tungsten	75 Re Rhenium	76 Os Osmium	77 Ir Iridium	78 Pt Platinum	79 Au Gold	80 Hg Mercury	81 Tl Thallium	82 Pb Lead	83 Bi Bismuth	84 Po Polonium	85 At Astatine	86 Rn Radon
7	87 Fr Francium	88 Ra Radium	89 Ac Actinium															
Lanthanides (Rare Earth Metals)					58 Ce Cerium	59 Pr Praseodymium	60 Nd Neodymium	61 Pm Promethium	62 Sm Samarium	63 Eu Europium	64 Gd Gadolinium	65 Tb Terbium	66 Dy Dysprosium	67 Ho Holmium	68 Er Erbium	69 Tm Thulium	70 Yb Ytterbium	71 Lu Lutetium
Actinides					90 Th Thorium	91 Pa Protactinium	92 U Uranium	93 Np Neptunium	94 Pu Plutonium	95 Am Americium	96 Cm Curium	97 Bk Berkelium	98 Cf Californium	99 Es Einsteinium	100 Fm Fermium	101 Md Mendelevium	102 No Nobelium	103 Lw Lawrencium

Reserved for New Artificially Produced Elements

FIG. 24–4 Periodic table of the elements. For atomic weights of the elements, see Table 21–1, p. 375.

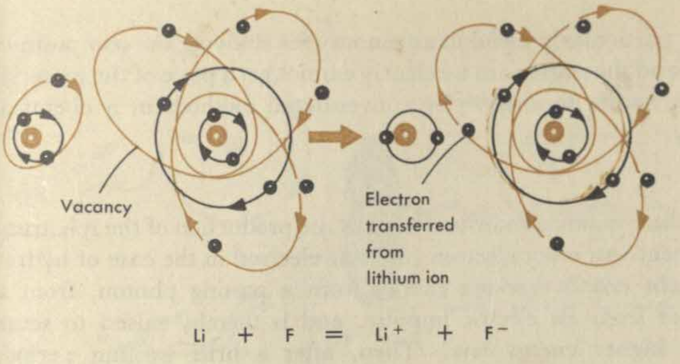


FIG. 24-5 The formation of lithium fluoride.

a sodium atom, it "adopts" the latter's lonely outer electron and becomes Cl^-, while the sodium atom becomes Na^+. The 2 ions are now held together by electrostatic forces and form a stable molecule of table salt. Similarly, an oxygen atom that has 2 electrons missing from its outer shell ($Z = 8 = 2 + 6$) tends to adopt 2 electrons from some other atom and can thus bind 2 monovalent atoms (H, Na, K, etc.) or 1 bivalent atom such as magnesium ($Z = 12 = 2 + 8 + 2$), which has 2 electrons to lend. An example of chemical binding of this kind is shown in Fig. 24-5.

24-5 Spectra of Multielectron Atoms

The single electron of hydrogen has only a relatively small number of energy levels, which are easily calculated by the Bohr theory. For this reason, it has a simple spectrum whose regularities were apparent. The 102 other and larger atoms have spectra which are much more complex and whose fundamental regularities are not at all apparent. Nevertheless, each element when vaporized has its own set of frequencies that it is able to emit or absorb, in a pattern as characteristic of that element as a set of fingerprints is of an individual human.

These characteristic spectra are produced by the outer electrons of the elements. But since each outer electron is influenced not only by the nucleus but by other electrons as well, the arrangement of possible energy levels can be very complicated, and the differences between energies (which control the frequencies emitted or absorbed) more complicated still.

The spectra of some elements (iron, for example) may have thousands of lines in the visible and near-visible range.

The fact that each chemical element emits a set of spectral lines characteristic only of that particular element is the basis of *spectral analysis*, the method by which we can observe the spectral lines emitted by a material of unknown chemical constitution and tell of what elements (and in what relative amounts) the substance is composed. Spectral

analysis is particularly useful in astronomy for studying the composition of the sun and the stars, since we clearly cannot get a piece of the material from these bodies to analyze by conventional methods in a chemical laboratory.

24-6 Lasers

So far, we have pictured only two steps in the production of the spectrum of an element. An outer electron (the *only* electron in the case of hydrogen) absorbs exactly enough energy from a passing photon, from a collision, or from an electric impulse, and is thereby raised to some permitted higher energy level. Then, after a brief waiting period, generally about 10^{-8} second, the electron *spontaneously* drops back to its lowest energy level, either in one jump or perhaps in several if there happen to be intermediate levels along the way, emitting one or more photons as it does so.

The only role we have assigned to the passing photon is to give up its entire energy in exciting the electron to a higher energy level. However, photons can and do play another part in this scheme. During the time the excited electron is pausing to wait for its spontaneous fall back down to a lower level, it is very susceptible to having this fall-back *stimulated*, or induced to take place sooner than otherwise, by interaction with a photon of the proper energy.

Imagine a population of excited atoms: in some the outer electron is in its lowest energy level E_1; some others are excited to the second level E_2; and some are in the E_3 state. Now consider a passing photon whose energy hf is equal to $E_3 - E_2$. We have already noted one effect it may have—it may interact with an electron in the E_2 state and give up its energy to raise the electron to E_3. But it may also interact with an electron in the E_3 state, to stimulate the electron to drop back at once and emit its $hf = E_3 - E_2$ photon. What makes the laser possible is that ***this emitted photon travels in exactly the same direction as the stimulating photon, and is exactly in phase with it.*** This is very different from a spontaneously emitted photon, which may be in any direction, and whose phase is not necessarily in step with that of any other photon.

Ordinarily this effect is not noticeable, because there are generally more atoms in the lower energy states, so the fate of most photons of the proper energy is to be absorbed in exciting the electrons. One of the tricks that makes a laser work is to provide more electrons in a state of higher energy. One type of laser uses a rod of synthetic ruby crystal—made of colorless aluminum oxide plus a fraction of a percent of chromium atoms. An intense flash of white light excites most of the chromium atoms to a higher-energy state which is *metastable*—that is, the electrons

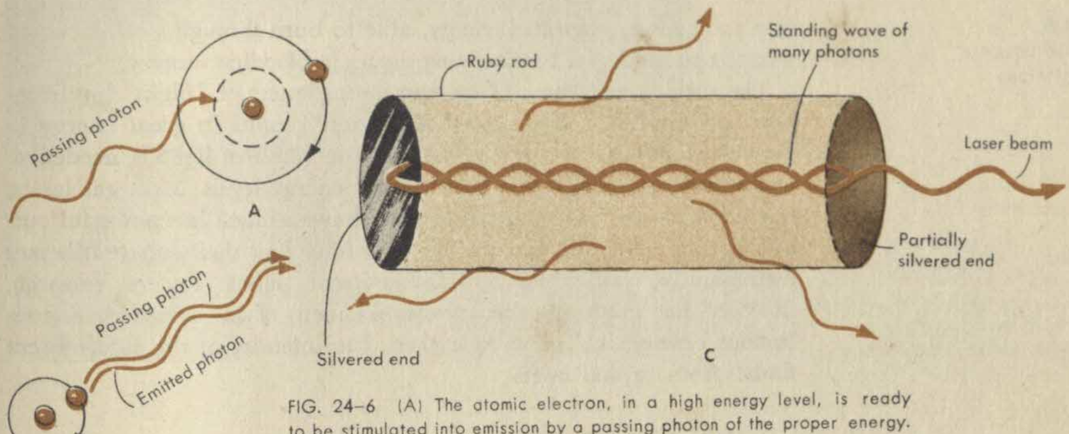


FIG. 24-6 (A) The atomic electron, in a high energy level, is ready to be stimulated into emission by a passing photon of the proper energy. (B) The electron has dropped back to a level of lower energy, emitting a photon exactly duplicating the stimulating photon. (C) Photons not parallel to the axis of the rod leak out of the cylindrical sides. The intense standing wave, composed of many coherent photons, can be started by one photon exactly in the axial direction.

remain at this level for the enormously long average time (atomically speaking) of about 10^{-3} second before spontaneously dropping back to a level of lower energy.

The earliest of these spontaneously produced photons stimulate the production of others; each one is exactly in phase with the stimulating photon, and traveling in exactly the same direction (Fig. 24-6B). If this were all, though, it would be only an interesting experiment, rather than an increasingly useful tool. To produce a laser of this type, the synthetic ruby is made in the shape of a cylinder, with its ends exactly flat and exactly parallel. These ends are now silvered—one completely, to reflect as perfectly as possible; and the other only partially, so that a fraction of the incident photons may escape.

An early spontaneous photon that is of just the right wavelength and is exactly perpendicular to the silvered ends, will be reflected back and forth between them to set up a standing wave (Fig. 24-6C). This stimulates the emission of other photons, which follow it exactly in both phase and direction, and these stimulate others, and so on until the standing wave grows to enormous intensity. Photons from this standing wave that leak through the partially silvered end constitute the actual laser beam.

This is a beam of *coherent* light. All the photons are in phase with each other, and all are traveling in the same direction. Such a coherent beam will spread very little. It can be focused by lenses or mirrors to a

tiny point of concentrated energy, able to burn through steel, to make microscopic welds, or to sear living tissues in bloodless surgery.

The ruby laser ("laser," from the initial letters of "***L***ight ***A***mplification by ***S***timulated ***E***mission of ***R***adiation") emits its great energy in very short pulses. Between pulses, another flash of light is needed to again raise the electrons to their higher energy levels. Most gas lasers, operating on the same basic principle, have a much less powerful output, but have the advantage of being able to emit their coherent beams continuously, rather than in intermittent pulses. Recent research, however, has made possible the development of gas lasers whose continuous beams rival, or even exceed, the intensity of the intermittent flashes from crystal lasers.

24-7 Continuous Spectra

Jeans's trouble with the "ultraviolet catastrophe" and Planck's solution for the difficulties arose from the analysis of radiation from hot bodies. If you examine the spectrum radiated from the hot tungsten filament of an incandescent lamp, you will find it to be continuous—*all* frequencies are present, and there are no lines to be seen. Iron vaporized in an electric arc emits a very definite pattern of many lines, by which the element iron can be identified. But a crucible of incandescent molten iron or a red-hot piece of solid iron will emit a continuous spectrum that cannot be distinguished from the emission from gold or tungsten or any other solid or liquid substance at the same temperature.

In a gas at relatively low pressure, the energy levels available to the outer electrons come from the atom itself, without disturbance from the outside. These energies have definite values, and the spectrum of billions of similar undisturbed atoms will include only certain definite frequencies. In solids or liquids, however, or even in compressed gases, the atoms are not alone and undisturbed. The energy levels are constantly and unpredictably being distorted by the electric fields of neighbors, who come and go and interfere continuously and randomly. Under such circumstances, all frequencies are possible and the spectrum is continuous.

The spectrum of the sun is an interesting combination. First accurately observed and measured by the German physicist Joseph Fraunhofer (1787–1826), the solar spectrum consists of a continuous spectrum crossed by many thousands of dark lines, each one indicating a frequency that is missing (or at least reduced in intensity) from the sun's radiation. The sun is composed entirely of hot gas, and its density rapidly increases with depth. In the deep and denser layers of the solar atmosphere, gas atoms are crowded so close together that they emit a continuous radiation

FIG. 24-7 The sun—a sphere of gas without a definite boundary. The photosphere is the layer of gas that emits the sunlight we see; some frequencies of this emission are absorbed in the chromosphere.

of all frequencies. The depth at which this continuous emission occurs is the *photosphere* (Fig. 24-7), which is what appears to us to be the "surface" of the sun as we observe it. In coming from the photosphere to our instruments, this radiation passes through thousands of miles of solar atmosphere of lower density, in which the atoms can act as individuals. In this less-dense region (the *chromosphere*), the atoms of each element present absorb their own characteristic frequencies. These absorbed frequencies we see as the dark *Fraunhofer lines*.

Analysis of these lines shows that, apart from the large predominance of hydrogen and helium gases, the chemical composition of the sun is identical with that of our earth. By using the methods of spectral analysis, astronomers have been able to learn a great deal, not only about the sun, but also about the chemical composition of the planetary atmospheres and the outer envelopes of distant stars.

24-8 X-Ray Spectra

Corresponding to the single electron of the hydrogen atom, the outermost electrons of other atoms also have orbits of higher energy to which they can jump when excited by heat or strong electrical fields or by the absorption of radiation of the proper frequency and energy. In dropping back to their normal levels, these electrons radiate frequencies characteristic of the atoms to which they belong. As in the hydrogen atom, these outer-electron frequencies are in the range of visible light, infrared, or ultraviolet. X rays are likewise produced by electron jumps from one energy level to another, but the energy differences radiated away are enormously greater than those associated with outer electrons, and the resulting X rays are of very high frequency.

An X-ray tube is similar in principle to a cathode-ray tube, and electrons emitted from the cathode are accelerated through a vacuum by

a potential of many thousands of volts, to strike against a metal anode. There are two ways in which this electron bombardment may cause high-frequency radiation to be emitted. First, the electron may be rapidly decelerated when it strikes the target. We know from Maxwell's work that any accelerated charge will emit electromagnetic radiation, and these electrons (since they are not in atomic orbits) are no exception. The deceleration of the free bombarding electrons is not quantized, so the resulting photons they emit can have any and all frequencies less than a certain limit. This limit comes from the principle of conservation of energy; the decelerating electron cannot emit a photon whose quantum of energy is greater than the kinetic energy of the electron. If, for example, the electrons are accelerated in the X-ray tube through a potential difference of 20,000 volts, each electron will have an energy of $E = Vq = 2 \times 10^4 \times 1.6 \times 10^{-19} = 3.2 \times 10^{-15}$ joule, or 3.2×10^{-8} erg. From Planck's $E = hf$, we see that the highest possible frequency that can be emitted is $f = 3.2 \times 10^{-8}/6.63 \times 10^{-27} = 4.8 \times 10^{18}$ hz. This corresponds to a wavelength of $\lambda = 3 \times 10^{10}/4.8 \times 10^{18} = 6.2 \times 10^{-9}$ cm, or 0.62 Å. This is shown as the "limiting wavelength" in Fig. 24-8.

Sometimes, however, a bombarding electron may actually knock an electron out of one of the inner shells of an atom of the target, or anode. When a heavy metal atom thus loses an inner electron, the vacancy is filled by one of the outer electrons falling down to take its place, and the energy differences between inner and outer electron shells are enormous, particularly for the atoms of heavy metals. These large energy differences, by the $E = hf$ relationship, are radiated away as energetic photons of very high frequency.

These photons of definite energy are shown as the spikes on the graph

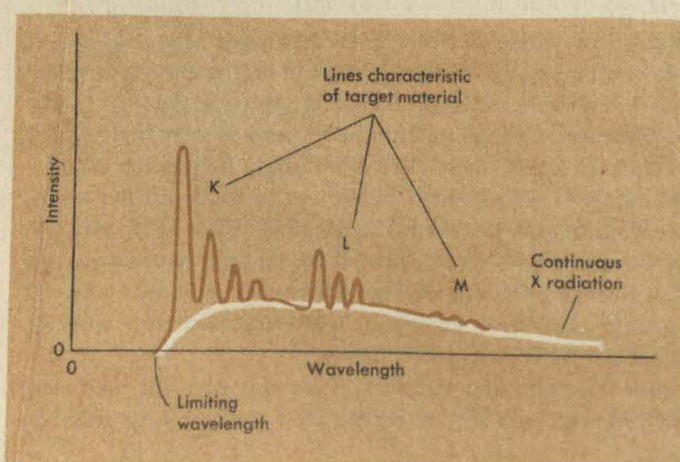


FIG. 24-8 Schematic diagram of an X-ray spectrum.

in Fig. 24-8. The series of lines marked K are caused by vacancies in the innermost K shell being filled by electrons from the L, M, and other outer shells. The lines marked L are emitted by electrons from outer shells falling into L shell vacancies, etc. Each element has its own particular schedule of energy differences, and thus each element has a characteristic X-ray spectrum by which it can be identified.

Questions

(24-1)

1. (a) For $n = 3$, what are the possible values for l? (b) For $l = 3$, what are the possible values for m?
2. (a) If $n = 4$, what values can l have? (b) What is the least possible value for l if there are shown to be five different values for m?
3. List in some tabular form each individual different set of four quantum numbers possible for $n = 2$.
4. In some systematic tabular form, list all the possible different sets of quantum numbers for $n = 3$.
5. The volume of a gram-atomic weight of iron (at. wt. 55.8, density 7.9 gm/cm^3) is $55.8/7.9 = 7.1$ cm^3. Since an atomic weight contains 6×10^{23} atoms, the volume of an iron atom is about $7.1/6 \times 10^{23} = 1.2 \times 10^{-23}$ cm^3. What is the approximate volume of an atom of (a) lithium, the lightest metal, at. wt. 6.94, density 0.53 gm/cm^3? (b) osmium, the densest metal, at. wt. 190, density 22.5 gm/cm^3?
6. Compare the approximate volumes of an atom of aluminum (at. wt. 27, density 2.7 gm/cm^3) and an atom of lead (at. wt. 207, density 11.4 gm/cm^3).
7. What are all the possible values for l for electrons in the N shell?

(24-2)

8. What are all the possible values for l for electrons in the M shell?
9. List the four quantum numbers for each of the electrons in an atom of beryllium.
10. List the four quantum numbers for each of the electrons in an atom of lithium.
11. How many electrons are in each occupied (or partially occupied) subshell in an atom of potassium?
12. How many electrons are in each occupied (or partially occupied) subshell in an atom of phosphorus?

(24-3)

13. In the series of "actinides," at the bottom of the periodic table, what subshell of what shell is being filled?
14. In the series of "lanthanides," what subshell of what shell is being filled?

(24-4)

15. Find, and justify by considering its atomic structure, the valence of (a) aluminum, (b) sulfur.
16. By considering its atomic structure, determine the valence of (a) calcium, (b) boron.
17. Sketch (similar to Fig. 24-5) the structure of a molecule of aluminum oxide.

18. Sketch (similar to Fig. 24-5) the structure of a molecule of boron chloride.

(24-7) **19.** In an X-ray tube, there is a potential difference of 12,500 volts between the cathode (which emits the electrons) and the anode (target). What is the wavelength of the highest frequency X rays that this tube can emit?

20. High-energy X rays of $\lambda = 0.08$ Å are to be produced by a special tube. What is the minimum potential difference that must be applied between its cathode and anode?

chapter / twenty-five

The Wave Nature of Particles

25-1 De Broglie Waves

Although Bohr's theory of atomic structure was successful in explaining a large number of known facts concerning atoms and their properties, the three fundamental postulates underlying his theory remained quite inexplicable for a long time. The first step in the understanding of the hidden meaning of Bohr's quantum orbits was made by a Frenchman, Louis de Broglie, who tried to draw an analogy between the sets of discrete energy levels that characterize the inner state of atoms and the discrete sets of mechanical vibrations that are observed in the case of violin strings, organ pipes, etc. "Could it not be," de Broglie asked himself, "that the optical properties of atoms are due to some kind of standing waves enclosed within themselves?" As a result of these considerations, de Broglie came out with his hypothesis that *the motion of electrons within the atom is associated with a peculiar kind of waves which he called "pilot waves."* According to this unconventional view, each electron circling around an atomic nucleus must be considered as being accompanied by a standing wave that runs around and around the electronic

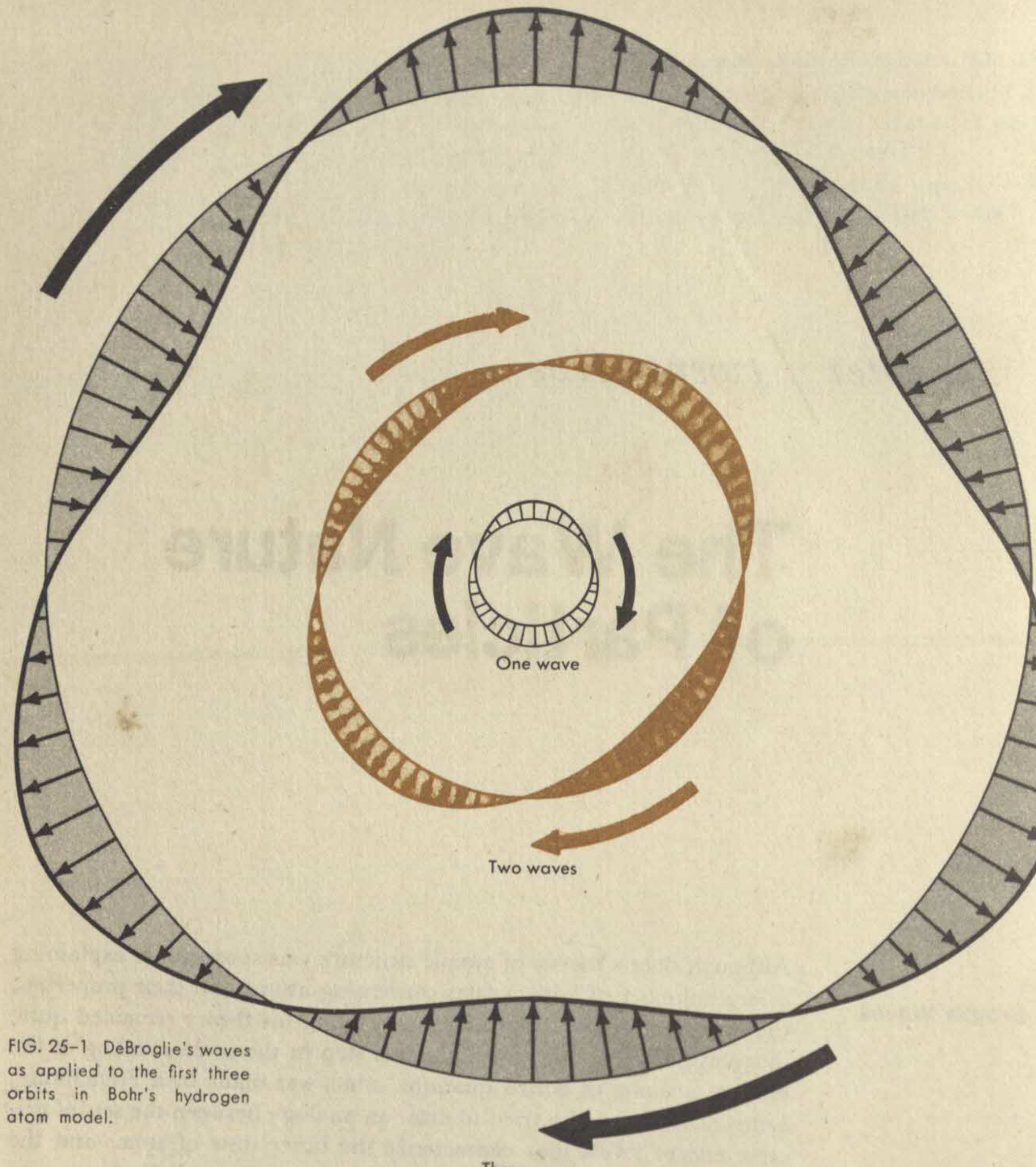


FIG. 25–1 DeBroglie's waves as applied to the first three orbits in Bohr's hydrogen atom model.

orbit. If this were true, the only possible orbits would be those whose lengths are an integral multiple of the wavelength of the corresponding *de Broglie wave.*

Figure 25-1 shows the application of de Broglie's idea to the first three orbits of the Bohr hydrogen atom. In order to have n complete wavelengths (λ_n) fit into the circumference of the nth orbit, the following relationship must be true:

$$n\lambda_n = 2\pi r_n.$$

In the development of Bohr's theory of the hydrogen atom, we had

$$r_n = \frac{n^2h^2}{4\pi^2me^2}.$$

Substituting this into the preceding equation, we get

$$n\lambda_n = \frac{n^2h^2}{2\pi me^2}$$

or

$$\lambda_n = \frac{nh^2}{2\pi me^2}.$$

The above expression for the length of the wave associated with an electron in the nth quantum orbit can be made simpler and more significant if we recall another relationship from the Bohr hydrogen atom—the velocity of the electron in the nth orbit:

$$v_n = \frac{2\pi e^2}{nh}.$$

Now, from our expression for λ_n, if we take out $nh/2\pi e^2$ and in its place substitute $1/v_n$, we shall have a much simpler expression:

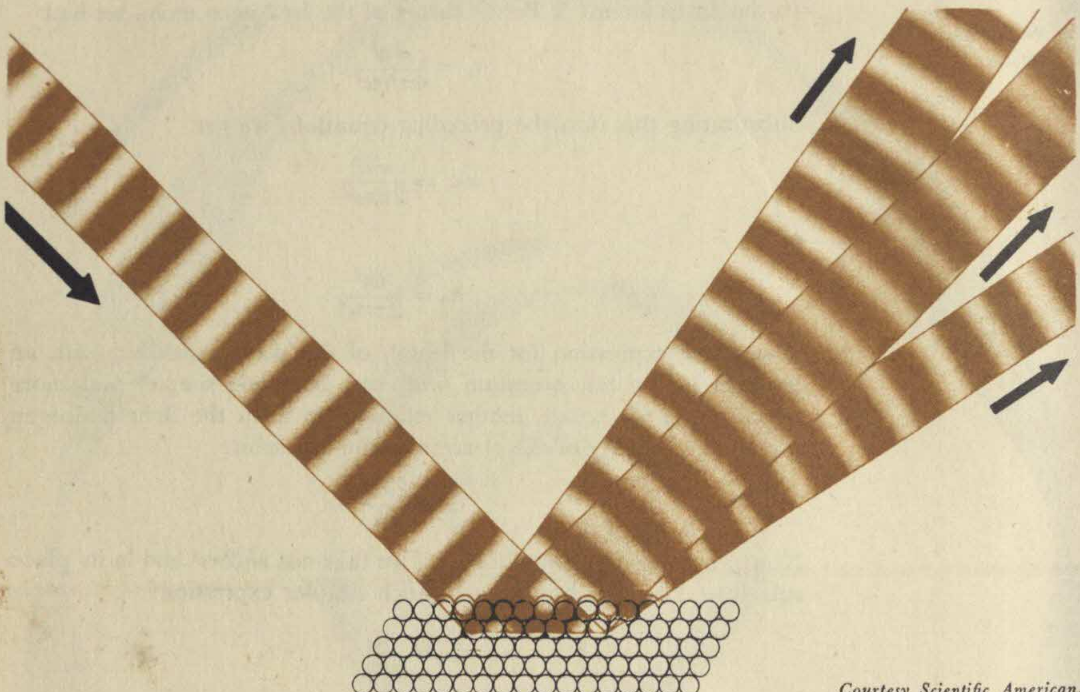
$$\lambda_n = \frac{h}{mv_n}.$$

The above relationship states de Broglie's fundamental hypothesis, that ***the wavelength of the wave associated with a moving particle is equal to Planck's quantum constant divided by the momentum of the particle.***

Although we have justified this by considering an electron in a Bohr atom, if electrons are really accompanied by these mysterious de Broglie waves while moving along the circular orbits within an atom, the same must be true for the free flight of electrons as observed in free electron beams. And, if the motion of electrons in the beams is associated with some kind of waves, we should be able to observe the phenomena of *interference* and *diffraction* of electron beams in the same way that we observe these phenomena in beams of light. For a stream of moving electrons, we assume for the wavelength of the waves the same formula that applied to orbital electrons within the atom:

$$\lambda = \frac{h}{mv}.$$

For the electron beams used in laboratories, this wavelength comes out to be much shorter than that of ordinary visible light and is comparable, in fact, with the wavelengths of X rays, i.e., about 10^{-8} cm. Thus it would be futile to try to observe the diffraction of electron beams by using ordinary optical diffraction gratings. We should instead employ a method similar to that used in studying X-ray spectra. To examine

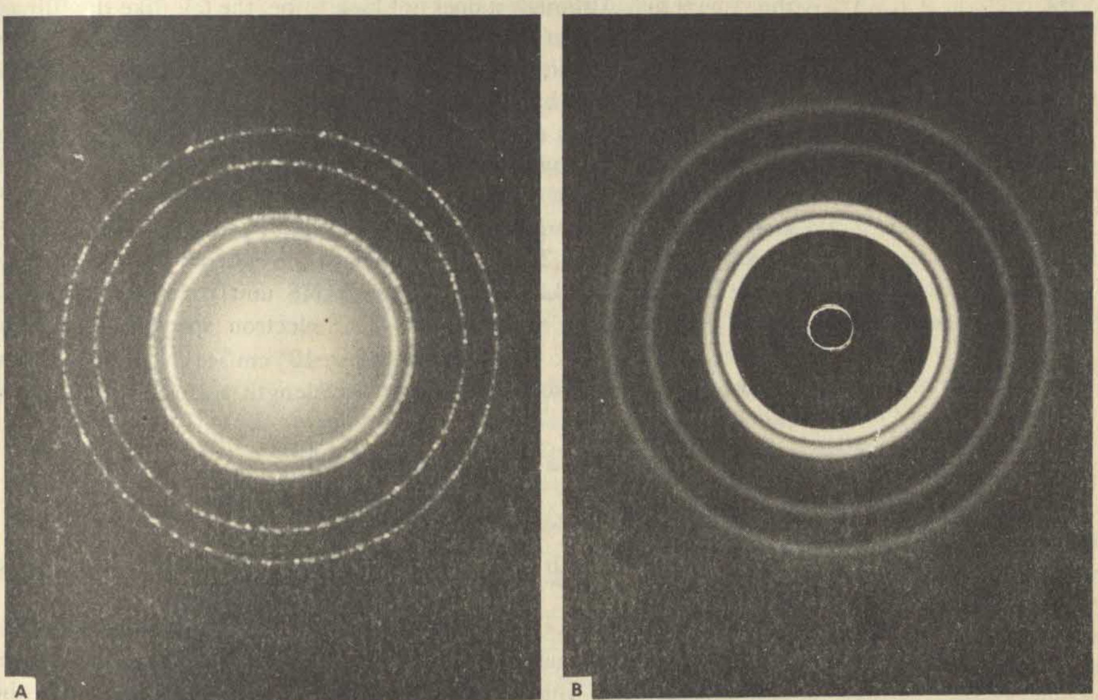


Courtesy Scientific American.

FIG. 25-2 A beam of X rays being reflected by the gratinglike molecular layers in a crystal lattice.

X-ray spectra, physicists use a "crystal spectrograph" that was developed by the British physicists W. H. and W. L. Bragg (father and son). The finest optical gratings that it is possible to make have a line spacing of several thousand Ångstroms, and are useless for studying the diffraction patterns of radiation whose wavelength is in the neighborhood of 1 Å. Nature, however, has provided suitable gratings. We have seen that the diameter of atoms is a few Ångstroms, and in the regular crystals of metals, salts, and other substances, the atoms are arranged in perfect rows and layers separated by this distance (Fig. 25-2). Incoming radiation is reflected, or scattered, by these regularly spaced crystal atoms, and the scattered Huygens wavelets are in phase only in certain directions. Their behavior is similar to that of light passing through a transmission grating as shown in Fig. 17-8. The resulting interference patterns are of course more complex because the crystal is equivalent to three gratings roughly at right angles to each other in three dimensions.

Figure 25-3A shows the pattern created by an X-ray beam passing through a layer of aluminum filings. The X rays, if they were diffracted by the regular array of atoms in a *single crystal*, would have produced



Photos from the Film Studio, Education Development Center.

FIG. 25–3 (A) Diffraction of a beam of X rays by a layer of aluminum filings. Wavelength = 0.71A. (B) Diffraction of a beam of electrons by thin aluminum foil. Wavelength = 0.50A. (The actual photographs have been enlarged by different amounts to make the resulting patterns the same size in this reproduction.)

a pattern of dots on the photographic film. The filings, being composed of a great number of crystals oriented in every possible direction, give the same pattern we would have if the single-crystal dots were rotated about the center line of the beam; all the dots the same distance from the center would combine to result in a circle.

Two American physicists, C. J. Davisson and L. H. Germer, used a similar arrangement in experiments on electron diffraction, the only difference being that the beam of X rays was replaced by a beam of electrons. The electrons were accelerated by a potential difference between a cathode and an anode, similar in principle to an X-ray tube. The beam was allowed to fall on the surface of a crystal, and the resulting interference pattern fully confirmed de Broglie's theory.

Figure 25-3B is the pattern produced by a beam of electrons passing through a piece of thin aluminum foil. Electrons are not penetrating enough to pass through a layer of filings, but aluminum foil accomplishes

the same thing. Although it does not look to be, the foil (like the filings) is made up of a great number of individual crystals oriented in every direction. The resulting electron-beam diffraction pattern is a duplicate of the one produced by the X rays.

As an example, let us consider a beam of electrons accelerated in an "electron gun" through a potential difference of 200 volts. Since 200 volts is 200 joules/coul and since the electron has a charge of 1.6×10^{-19} coul, each electron gains an energy $E = Vq = 200 \times 1.6 \times 10^{-19} = 3.2 \times 10^{-17}$ joule. This is in the form of the electron's kinetic energy, so we have $\frac{1}{2}mv^2 = 3.2 \times 10^{-17}$ joule (MKS unit!), and $v^2 = 2 \times 3.2 \times 10^{-17}/9.1 \times 10^{-31} = 7.0 \times 10^{13}$. The electron speed is then $v = \sqrt{70 \times 10^{12}} = 8.4 \times 10^6$ m/sec or 8.4×10^8 cm/sec. Now we can use de Broglie's relationship to find the wavelength associated with these moving electrons:

$$\lambda = \frac{h}{mv} = \frac{6.63 \times 10^{-27}}{9.1 \times 10^{-28} \times 8.4 \times 10^8}$$

$$= 8.6 \times 10^{-9}\ \text{cm}, \quad \text{or} \quad 0.86\ \text{Å}.$$

This is a wavelength very suitable for showing diffraction by most crystals.

A few years later, a German physicist, Otto Stern, repeated the experiments of Davisson and Germer by using a molecular beam of sodium atoms (Fig. 25-4) instead of an electron beam and found that the

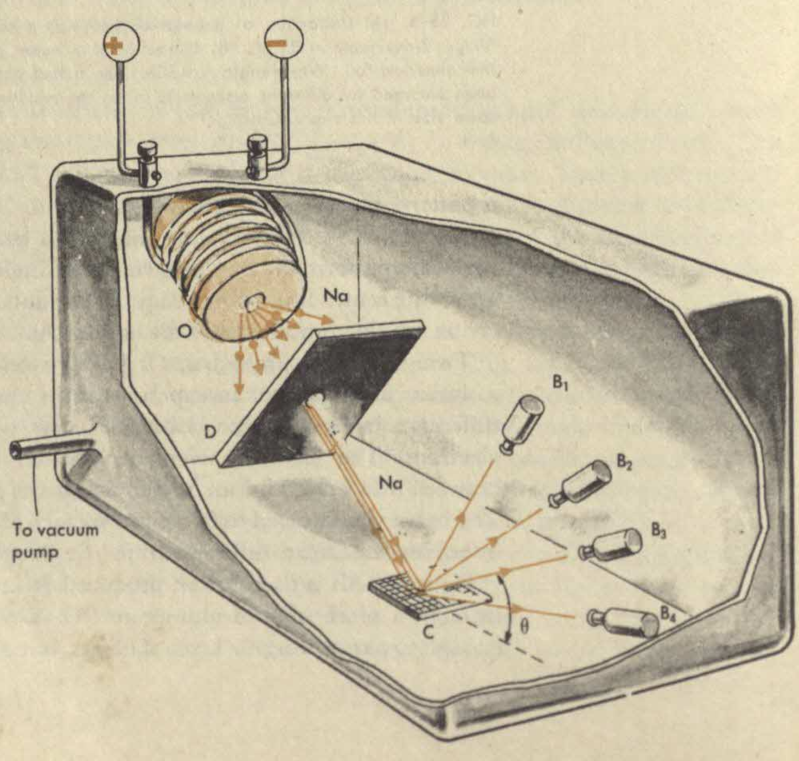


FIG. 25–4 O. Stern's demonstration of the diffraction of atoms. Some of the atoms from the oven O pass through the small hole in the barrier *D*, from which they emerge as a narrow beam. The atoms reflected from the crystal are trapped and measured in the collectors *B*. The results show a strong maximum in the direction required by the ordinary law of reflection, plus a number of other secondary maxima corresponding to a diffraction pattern.

diffraction phenomenon exists in that case, too. Thus it became quite evident that, in material particles as small as atoms and electrons, the basic ideas of classical Newtonian mechanics should be radically changed by introducing the notion of waves associated with material particles in their motion.

The once definite distinction between waves and particles seems to have broken down. There are many sorts of interference experiments in which light shows itself to be unquestionably a wave phenomenon; yet in the photoelectric effect it concentrates all its energy on a single electron, as though it were a bulletlike particle. And now electrons and atoms, so surely particles, behave in some experiments as though they were waves.

The wave aspect and the particle aspect seem to be so mutually contradictory that it is quite natural to ask which one is "really" correct for a beam of light or an electron. The modern physicist will say that neither one is "really" correct. We are trying to make the submicroscopic world of the photon and the atom fit models we imagine as being analogous to tiny bullets and tiny ripples on water tanks. The world of the atom and the photon cannot be described in the same terms we use to describe the behavior of the macroscopic world of matter-in-bulk. That we cannot is attested to by the wave-particle dilemma and the contradictions of a similar nature we run into when we try. The physicist has mathematical equations whose solutions will give the correct answers, whether it be wave or particle that is involved; but to these equations no model or picture is connected. We must learn either not to ask questions about which model is "really" correct or not to think of any model at all. We quite plainly cannot have it both ways. An intricate mathematical method for handling this kind of problem was worked out by an Austrian physicist, E. Schrödinger, and it represents the subject matter of an important but rather difficult branch of modern theoretical physics known as *wave mechanics*, or *quantum mechanics*.

25-2 The Uncertainty Principle

Now things seem to be going from bad to worse. First, we had Bohr's "quantized orbits" that looked like railroad tracks along which the electrons were running around the atomic nucleus. Then these tracks were replaced by mysterious "pilot waves" that were supposed to provide "guidance" for the electrons in their orbital motion. It all seemed to be against common sense, but, on the other hand, these developments of the quantum theory provided us with the most exact and most detailed explanation (or description) of the properties of atoms—their spectra, their magnetic fields, their chemical affinities, etc. How could it be? How could such a picture, nonsensical at first sight, lead to so many positive results? Here we have to repeat what was already said in con-

nection with Einstein's theory of relativity. Modern physics extends its horizons far beyond the everyday experience upon which all the "common sense" ideas of classical physics were based, and we are thus bound to find striking deviations from our conventional way of thinking and must be prepared to encounter facts that sound quite paradoxical to our ordinary common sense. In the case of theory of relativity, the revolution of thought was brought about by the realization that space and time are not the independent entities they were always believed to be but are the parts of a unified space-time continuum. In the quantum theory, we encounter the nonconventional concept of *a minimum amount of energy*, which, although of no importance in the large-scale phenomena of everyday life, leads to revolutionary changes in our basic ideas concerning motion in tiny atomic mechanisms.

Let us start with a very simple example. Suppose we want to measure the temperature of a cup of coffee, but all we have is a large thermometer hanging on the wall. Clearly, the thermometer will be inadequate for our purpose because when we put it into the cup it will take so much heat from the coffee that the temperature shown will be considerably less than that which we want to measure. We can get a much better result if we use a small thermometer that will show the temperature of the coffee and take only a very small fraction of its heat content. The smaller the thermometer we use for this measurement, the smaller is the disturbance caused by the measurement. In the limiting case, when the thermometer is "infinitely small," the temperature of the coffee in the cup will not be affected at all by the fact that the measurement was carried out. The common sense concept of classical physics was that this is always the case in whatever physical measurements we are carrying out, so that we can always compute the disturbing effect of whatever gadget is used for the measurement of some physical quantity and get the exact value we want. This statement certainly applies to all large-scale measurements carried out in any scientific or engineering laboratory, but it fails when we try to stretch it to such tiny mechanical systems as the electrons revolving around the nucleus of the atom. Since, according to Max Planck and his followers, energy has "atomic structure," *we cannot reduce the amount of energy involved in the measurement below one quantum*, and making exact measurements of the motion of electrons within an atom is just as impossible as measuring the temperature of a demitasse of coffee by using a bulky bathtub thermometer! But, whereas we can always get a smaller thermometer, it is absolutely impossible to get less than one quantum of energy.

A detailed analysis of the situation indicates that the existence of minimum portions (quanta) of energy prevents our being able to describe the motion of atomic particles in the conventional way, by calculating their successive positions and velocities. Quantum mechanics

shows us that the position of a particle and its momentum can be known only within certain limits. These limits are negligibly small for large-scale objects, but become of great importance in the submicroscopic world of atoms and atomic particles.

The uncertainty in our knowledge of the coordinate x and the momentum p of any particle can be expressed by writing $x \pm \Delta x$ and $p \pm \Delta p$, which means that all we can say is that the particle is located somewhere between $x - \Delta x$ and $x + \Delta x$, and the momentum of the particle lies somewhere between $p - \Delta p$ and $p + \Delta p$. The German theoretical physicist Werner Heisenberg showed in 1927 that these uncertainties are related by the expression

$$\Delta x \times \Delta p = \frac{h}{2\pi},$$

or

$$\Delta x \times m\Delta v = \frac{h}{2\pi}.$$

If we apply Heisenberg's uncertainty principle to a particle with a mass of 1 mg, we have that

$$\Delta x \times \Delta v = \frac{h}{2\pi m} \approx 10^{-24},$$

which tells us that if we know the position of the particle to within $\pm 10^{-12}$ cm, we are permitted by Nature to know the particle's velocity only to within $\pm 10^{-12}$ cm/sec. Clearly, such small uncertainties are of no significance when we are dealing with milligram-sized particles or larger. However, for an electron ($m \approx 10^{-27}$ gm) we have that

$$\Delta x \times \Delta v \approx 1,$$

so that if we have an uncertainty of 10^{-6} cm in the location of an electron, we cannot know its velocity with any greater precision than $\pm 10^{6}$ cm/sec. These uncertainties are large enough to make the classical picture of definite particles in definite orbital motion completely invalid. De Broglie's waves give us a new way of describing the behavior of atomic "particles."

25-3 Waves of Probability

When light was shown to have wave properties, the nineteenth-century physicists at once asked, "waves in what?" In order to answer this question, they invented the "ether," with all of its contradictory properties, only to have it taken away from them by Einstein.

It was not necessary to invent a medium for the de Broglie waves, but it was necessary to invent an *interpretation* of them. It turns out that the waves have no real material characteristics but are purely a measure of *probability*. And since the waves are an inescapable attribute of every

material particle, we must regretfully conclude that we can never fully know the momentum of a particle, or its location, but can at best only say that some certain momentum is more probable than another or that the particle is more likely to be here than there.

Fundamentally, this was the basic difficulty with Bohr's atomic model. It pictured an atomic electron as a particle with a definite location, and a definite velocity and momentum at all times. If a modern physicist were forced to make some sort of a model, he might picture the electron of a hydrogen atom, say, as a pattern of standing waves. These standing waves are not confined by definite boundaries, as are those of a vibrating violin string or the electromagnetic waves in Jeans's imaginary reflecting cube. They are instead confined by the inverse-square Coulomb force of attraction between the nucleus and the electron's negative charge. Since this force extends out more and more weakly without limit, there is a *very* small chance that any given electron may be anywhere in the universe. However, calculations show that in the hydrogen atom in its lowest-energy $n = 1$ state, the probability waves, which can be pictured as spreading out in a sort of cloud of varying density, are most dense at a distance of 0.53 Å from the nucleus. This is where the wave mechanics calculations say the electron is most likely to be, and this is exactly the radius of Bohr's first hydrogen orbit!

In spite of this agreement, there is a great conceptual difference between the Bohr model and the quantum mechanics interpretation. In the Bohr model, there was no real reason for saying that angular momentum must come in discrete units of $h/2\pi$, except that this is what *must* be assumed if the answers are to come out right and in agreement with the experimental spectral data. In the solutions of Schrödinger's probability-wave equation, no assumptions of this sort are necessary, and the quantum numbers n, l, and m come out with their allowed integral values as a part of the answer. The probability cloud takes on different shapes and dimensions to represent the higher-energy excited states, but still remains a diffuse cloud.

If the shape and density of this cloud is all we can ever know about an electron (or, if indeed, this is all there *is* to know about it), then the uncertainty principle follows naturally as an inescapable, built-in characteristic of how the world is made.

Questions

(25-1)

1. What is the wavelength of an electron moving with a velocity of 2×10^7 cm/sec?

2. The speed of an electron is 1.2×10^6 m/sec. What is its de Broglie wavelength?

3. Interference effects show that the wavelength associated with a beam of electrons is 3.2×10^{-8} cm. What is the velocity of the electrons in the beam?

4. With what speed must an electron travel in order to have a de Broglie wavelength of 10^{-8} cm?

5. What is the wavelength associated with the hydrogen electron in its $n = 2$ orbit?

6. What wavelength is associated with a hydrogen electron in its $n = 3$ orbit?

7. (a) What is the kinetic energy of the electrons in Question 3? (b) Through what potential difference must the electrons have been accelerated in order to have this much energy?

8. (a) What is the kinetic energy of an electron that has a de Broglie wavelength of 10^{-8} cm? (b) Through what potential difference must it have been accelerated?

9. A proton and an electron are given the same kinetic energy ($m_p = 1840m_e$). (a) How do their speeds compare? (b) How do their momenta compare? (c) How do their de Broglie wavelengths compare?

10. A proton and an alpha particle are found to have the same kinetic energy ($m_\alpha = 4m_p$). (a) What is the ratio of their speeds? (b) What is the ratio of their momenta? (c) What is the ratio of their de Broglie wavelengths?

11. What wavelength is associated with a beam of electrons accelerated through a potential difference of 75 volts?

12. What wavelength is associated with a beam of protons (See Question 9) accelerated through a potential difference of 75 volts?

13. Vaporized sodium (at. wt. = 23) has a temperature of about 1165°K, which gives the atoms of vapor an *rms* speed of about 700 m/sec. What is the de Broglie wavelength of a sodium atom at this speed?

14. Potassium (at. wt. = 39.1) vaporizes at a temperature of 1047°K, and the atoms of potassium vapor at this temperature have an *rms* speed of about 510 m/sec. What is the wavelength of a potassium atom at this speed?

(25-2) **15.** Take one of the droplets in Millikan's oil-drop experiment and assume it to have a mass of 10^{-12} gm. If it were possible to determine its position at any moment with an accuracy of $\pm 10^{-3}$ cm, what must be the irreducible minimum of uncertainty as to its speed at that moment?

16. A 1-gm bullet is fired with a speed of 300 m/sec, which the experimenter knows is accurately determined to within 0.01 percent. What restriction does the uncertainty principle put on the determination of the bullet's position at any time?

17. Electrons in a beam have a speed of 5×10^7 cm/sec, with an uncertainty of 1 part in 10^3. What is the least possible uncertainty the experimenter may have in the position of any electron?

18. A beam of protons (See Question 9) has a speed of 5×10^7 cm/sec, with an uncertainty of 0.1 percent. What is the least possible uncertainty in the location of a proton?

19. An electron is (in imagination) confined in a cubical box 1 mm on a side. We therefore know its location to within 0.5 millimeter. (a) What is our mini-

mum uncertainty about its speed? (b) Can the electron be definitely at rest within the box? Explain.

20. In the Bohr model of the hydrogen atom, we can never know just where an electron is in its orbit; our uncertainty of its position is therefore $\pm r$. Nor can we know whether at any instant it is moving (say) upward with a momentum of mv, or downward with a momentum of $-mv$; our uncertainty of its momentum is thus $\pm mv$. Multiply together the Bohr formulas for these uncertainties, and check on whether Heisenberg is vindicated by this interpretation.

chapter / twenty-six

Radioactivity and the Nucleus

26-1 Discovery of Radioactivity

The discovery of radioactivity, like that of many other unsuspected aspects of physics, was purely accidental. It was discovered in 1896 by the French physicist A. H. Becquerel (1852–1908), who was interested at that time in the phenomenon of fluorescence, which is the ability of certain substances to transform ultraviolet radiation that falls on them into visible light. In one of the drawers of his desk, Becquerel kept a collection of various minerals that he was going to use for his studies, but, because of other pressing matters, the collection remained untouched for a considerable period of time. It happened that in the drawer there were also several unopened boxes of photographic plates, and one day Becquerel used one of the boxes in order to photograph something or other. When he developed the plates, he was disappointed to find that they were badly fogged as if previously exposed to light. A check on the other boxes in the drawer showed that they were in the same poor condition, which was difficult to understand, since all the boxes were sealed and the plates inside were wrapped in thick black paper. What could

have been the cause of this mishap? Could it have something to do with one of the minerals in the drawer? Being of inquisitive mind, Becquerel investigated the situation and was able to trace the guilt to a piece of uranium ore labeled "Pitchblende from Bohemia." One must take into account, of course, that at that time uranium was not in vogue as it is today. In fact, only a very few people had ever heard about this comparatively rare and not very useful chemical element.

But the ability of uranium ore to fog photographic plates through a thick cardboard box and a layer of black paper rapidly brought this obscure element to a prominent position in physics.

The existence of penetrating radiation that can pass through layers of ordinarily opaque materials, as if they were made of clear glass, was a recognized fact at the time of Becquerel's discovery. In fact, only a year earlier, in 1895, a German physicist, Wilhelm Roentgen (1845–1923), discovered, also by sheer accident, what are now known as X rays, which could penetrate equally well through cardboard, black paper, or the human body. But, whereas X rays could be produced only by means of special high-voltage equipment shooting high-speed electrons at metallic targets, the radiation discovered by Becquerel was flowing steadily, without any external energy supply, from a piece of uranium ore resting in his desk. What could be the origin of this unusual radiation? Why was it specifically associated with the element uranium and, as was found by further studies, with two other heavy elements, thorium and actinium?

The early studies of the newly discovered phenomenon, which was called *radioactivity*, showed that the emission of this mysterious radiation was completely unaffected by physical and chemical conditions. One can put a radioactive element into a hot flame or dip it into liquid air without the slightest effect on the intensity of the radiation it emits. No matter whether we have pure metallic uranium or its chemical compounds, the radiation flows out at a rate proportional to the amount of uranium in the sample. These facts led the early investigators to the conclusion that the phenomenon of radioactivity is so deeply rooted in the interior of the atom that it is completely insensitive to any physical or chemical conditions to which the atom is subjected.

Becquerel's discovery attracted the attention of Polish-born Madame Marie Sklodowska Curie (1867–1934), wife of the French physicist Pierre Curie. She suspected that the radioactivity of uranium ore might, to a large extent, be due to some other chemical element, much more active than uranium, which might, however, be present in uranium ore in very small quantity. Being an experienced and hard-working chemist, Madame Curie decided to separate this hypothetical element from uranium ore by a painstaking method known as "chemical

fractioning." Loads of pitchblende from Bohemia went through Madame Curie's chemical kitchen, where careful processing was taking place, and only the fractions of the material emitting strong radiation were retained. Her work culminated in 1898 with a brilliant success: she obtained about 200 mg of a pure element that was a million times more radioactive than uranium itself. She christened the new element *radium* and established its number in the periodic system of elements as 88 ($Z = 88$). Another radioactive element discovered by Madame Curie was even more active than radium, and she named it *polonium* in honor of her native country. The study of radioactivity carried on by many investigators at the turn of the century led to the discovery of many other radioactive elements which early experimenters gave such strange names as "uranium X_1," "ionium," "radium emanation," etc.

26-2 The Nature of the Nucleus

It is obvious to us now that the mysterious radiation emitted by certain elements of high atomic number is an activity of the atomic nucleus. Madame Curie and her contemporaries, however, did not have the advantage of our more modern knowledge. (Remember that her separation of polonium and radium took place 13 years before Rutherford's scattering experiments first showed that atoms even have a nucleus!)

Rather than grope with our ancestors through decades of speculation until the neutron was discovered in 1932, we shall help our understanding more by taking advantage of our superior historical position. In a later chapter, we shall look at the nucleus and its behavior in more detail; for our present purpose, it will be enough to merely sketch some of its main features.

In the Bohr atom, we identified the hydrogen nucleus to be a single *proton*—a particle having a mass of 1.673×10^{-24} gm, about 1836 times the mass of an electron. Its charge is the same as the charge on an electron (1.60×10^{-19} coul, or 4.80×10^{-10} esu), but is positive.

But what of the nuclei of the heavier elements? It was suspected in the early days that they were made up primarily of aggregations of protons. Oxygen, for example, with a mass about 16 times the mass of a hydrogen atom, was thought to have a nucleus of 16 protons. Since its positive charge is only $8e$, it was assumed that the oxygen nucleus also contained 8 electrons to neutralize the proper part of the proton charge. Quantum mechanics, however, showed that the de Broglie waves of these supposed nuclear electrons could not possibly fit in the small volume of a nucleus. As we shall see, the discovery of the *neutron* in 1932 solved this uncomfortable problem. The neutron has a mass slightly greater than the proton (1.675×10^{-24} gm) and has no electrical charge.

Although there is still a great deal that we do not understand about

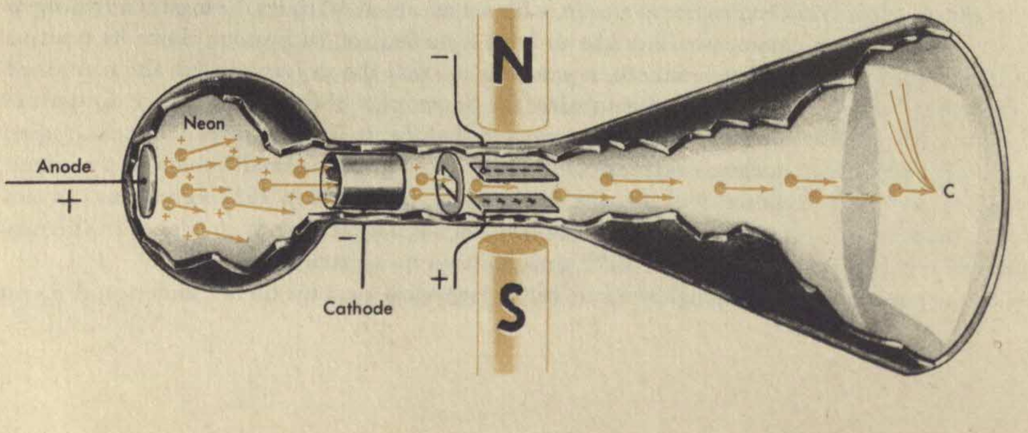
nuclei, we do know that for most practical purposes we can consider them to be aggregations of protons and neutrons. The oxygen nucleus, with a charge of $+8e$, must contain 8 protons; to make up the rest of the required mass, it also contains 8 neutrons. Armed with this basic knowledge, we are in a better position to understand much of the early experimental work.

26-3 Canal Rays and Isotopes

While the study of cathode rays in Thomson's tube led to the discovery of electrons, the study of *canal rays*, which are streams of positively charged gas ions, was also very helpful to the understanding of the inner nature of the atom. The apparatus Thomson used for the study of canal rays was a modification of the tube used for determining the e/m ratio and is shown in Fig. 26-1. A small amount of gas is left within the tube, and when a swiftly moving electron collides with a gas molecule, an electron is likely to be knocked off, making the molecule into an ion with a positive charge. The mass and electric charge of these positively charged canal rays can be analyzed by deflecting them in parallel electric and magnetic fields.

In his e/m work, Thomson used electric and magnetic fields perpendicular to each other, so that the two deflections could be made to add up to zero. If you go back and check over that section, you will see that his approach was possible only because all the electrons in his beam had the same velocity. This was because all of them originated on the

FIG. 26-1 The apparatus that led to the discovery of isotopes. Positive neon ions were accelerated by an electric field and formed a thin beam after passing through a slit. The beam was deflected by an electric field and a magnetic field, and fell on a fluorescent screen at the end of the tube. If all of the neon ions had the same mass, the trace on the screen would have been a single parabola (different points on the parabola corresponding to different velocities). Actually, there were three different parabolas, corresponding to masses of 20, 21, and 22.



cathode, and all were accelerated through the same total potential difference between cathode and anode.

With the canal rays, composed of positive ions, the situation was different. Neutral neon atoms were of course not affected by the electric field between cathode and anode. Their ionization by the energetic electrons streaming out from the cathode was caused by collisions that could take place almost anywhere in the field. The result was that some positive ions would be accelerated through only a small part of the field, some through nearly all of it, and others in between. The beam of canal rays therefore consisted of a stream of positive neon ions with a wide range of speeds.

Although we shall not stop to do it, it is not difficult to show that such a stream of ions, all of the same charge and mass but with different velocities, will leave a trace in the shape of a parabola. The mass of the ions can be computed quite accurately from the geometry of the parabola.

In measuring the mass of the particles forming canal rays in a tube filled with neon gas, Thomson expected to confirm the chemical value of the atomic weight of neon, which was known to be 20.182. However, instead of this value he got only 20.0, which was considerably lower, and well beyond the limits of possible experimental error. The discrepancy was explained when Thomson noticed that the beam of neon ions passing through the magnetic and electric fields was not deflected as a single beam, but was split into three branches, as shown in Fig. 26-1. (Actually, the third branch was discovered later by Thomson's co-worker, F. W. Aston, using improved apparatus.) The particles in the main branch, containing over 90.5 percent of all the neon ions, had a mass value of 20.0; the other fainter branch contained about 9.2 percent, and had a mass of 22.0. And a still fainter branch contained 0.3 percent of mass 21.

This was very remarkable! Here Thomson had found three kinds of neon atoms, *identical in chemical nature and having the same optical spectra, but different in mass.* On top of this, the mass values were almost exactly integral numbers. Ordinary neon, then, was actually a mixture of three different neons, and the chemical weight was just the average weight of this mixture.

The different types of neon were called *isotopes* of this element, which means in Greek "same place" and refers to the fact that all the neons of different weight occupy the same place in the table of elements. We can confirm this average weight by taking the weighted average from Thomson's data:

$$\begin{aligned} 0.905 \times 20 &= 18.10 \\ 0.092 \times 22 &= 2.02 \\ 0.003 \times 21 &= \underline{0.06} \\ \text{Average mass} &= 20.18. \end{aligned}$$

We can understand now just what the difference is between the three isotopes of neon. Since the atomic number of neon is 10 ($Z = 10$), we know that each atom must have 10 protons in its nucleus. If this number were *not* 10, it would have a different number of atomic electrons and would be a different element, rather than neon. To make up the observed masses, the isotope of mass 20 must have $20 - 10 = 10$ neutrons in its nucleus; the neon-21 must have $21 - 10 = 11$ neutrons; and the neon-22, 12 neutrons.

There is a fairly standardized system of notation for indicating different isotopes. In front of the letters of the chemical symbol of the element, a subscript gives the *atomic number* (Z), which is the number of protons in the nucleus. After the symbol, the *mass number* is placed as a superscript. The mass number is merely the total number of protons plus neutrons in the nucleus—or, more simply, the total number of *nucleons*, a term used to include both protons and neutrons. Thus, for our three isotopes of neon, we could write ${}_{10}Ne^{20}$, ${}_{10}Ne^{21}$, and ${}_{10}Ne^{22}$. (The prefix subscript 10 is not necessary; if the element is neon, this number *must* be 10, and it is therefore redundant. Nevertheless, for convenience, it is often used.)

Thomson's original crude apparatus has been improved by Aston, A. J. Dempster, and K. T. Bainbridge; and the modern *mass spectrograph* can determine the relative masses of isotopes with great accuracy.

Further studies have shown that almost every element represents a mixture of several isotopes. While in some cases (as in gold and iodine), one isotope accounts for 100 percent of the material, in many other cases (as in chlorine and zinc), different isotopes have comparable abundances. The isotopic composition of some of the chemical elements is shown in Table 26-1.

TABLE 26-1 ISOTOPIC COMPOSITION OF SOME ELEMENTS

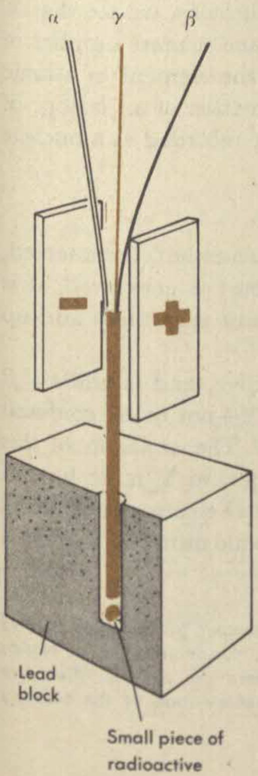
Atomic Number	Name	Isotopic Composition Percentage Shown in Parentheses
1	Hydrogen	1 (99.985); 2 (0.015)
6	Carbon	12 (98.9); 13 (1.1)
7	Nitrogen	14 (99.64); 15 (0.36)
8	Oxygen	16 (99.76); 17 (0.04); 18 (0.20)
17	Chlorine	35 (75.4); 37 (24.6)
30	Zinc	64 (48.89); 66 (27.81); 67 (4.07); 68 (18.61); 70 (0.62)
48	Cadmium	106 (1.215); 108 (0.875); 110 (12.39); 111 (12.75); 112 (24.07); 113 (12.26); 114 (28.86); 116 (7.58)
80	Mercury	196 (0.15); 198 (10.02); 199 (16.84); 200 (23.13); 201 (13.21); 202 (28.80); 204 (6.85).

26-4 Alpha, Beta, and Gamma Rays

Armed with a little twentieth-century knowledge, we can better understand some of the problems uncovered by the research of earlier years. Becquerel and his followers were puzzled to find that the radiation from their impure mixtures of radioactive elements consisted of three different components.

It is quite easy (in principle, at least) to separate these three types of radiation when they are emitted from a small piece of material containing a mixture of radioactive elements. If we drill a small hole in a block of lead, which is a good absorber of radiations of all kinds, and place a speck of radioactive material at the bottom of the hole, a narrow, well-defined beam of radiation will be emitted from the top of the hole (Fig. 26-2). If this beam is passed through a strong electric field between a pair of parallel plates, the single beam will be separated into its three components, as shown. The same separation will follow, if, instead of an electric field, the beam is passed through a strong magnetic field perpendicular to the plane of the drawing. Lacking knowledge of just what these radiations were, the experimenters named them simply alpha (α), beta (β), and gamma (γ) radiation, from the first three letters of the Greek alphabet.

FIG. 26-2 The separation of alpha, beta, and gamma radiation by passage through an electric field. (The same sort of separation can also be accomplished by a magnetic field.)



From their behavior in electric and magnetic fields, it was apparent that the α radiation consisted of positively charged particles; the β radiation of negatively charged particles; and the γ radiation was either neutral particles or electromagnetic radiation. They were all soon identified.

Alpha rays were found (by Rutherford) to consist of the fast-moving nuclei of helium atoms, now called α particles. From the atomic weight and atomic number of helium, we see that the helium nucleus must contain 2 protons and $4 - 2 = 2$ neutrons, making 4 nucleons altogether. Thus, when an unstable nucleus ejects an α particle, it loses 4 nucleons, and its *mass number is reduced by 4.* And, since 2 of these nucleons are protons, its *atomic number is reduced by 2,* and it of course then becomes another element!

Beta rays were found to be nothing more than energetic, fast-moving electrons. No one has ever been able to detect any difference between these electrons emitted from nuclei and the ordinary atomic electrons with which we are already familiar. But we have said that an electron cannot exist inside a nucleus; how, then, can a nucleus emit something it did not contain in the first place? The answer to this is that the electron is manufactured as it is emitted. The mass of an electron, 9.11×10^{-28} gm, can be produced (via Einstein's $E = mc^2$) from $9.11 \times 10^{-28} \times 9 \times 10^{20} = 8.20 \times 10^{-7}$ erg of energy. To account for the electron's negative charge, we may think of the neutral neutron as containing both $+e$ and $-e$ of charge. If the $-e$ is carried away by the electron, then the nuclear particle that was a neutron becomes a nucleon with a charge of $+e$—in other words, a neutron has been changed to a proton.

Thus, when a β particle is emitted, the *mass number is unchanged* and the *atomic number is increased by 1.*

Gamma rays are the pure energy of electromagnetic radiation and are emitted together with some α particles and some β particles. They do not change the number of nucleons, and they have no charge. Therefore, *the mass number and the atomic number are both unchanged.*

26-5 Families of Radioactive Elements

Radioactivity, as observed by Becquerel, turned out to be a composite effect that was due to the presence of a large number of radioactive elements. In fact, studies by the British physicist Soddy and his famous collaborator Rutherford showed that this mixture contained over a dozen individual elements.

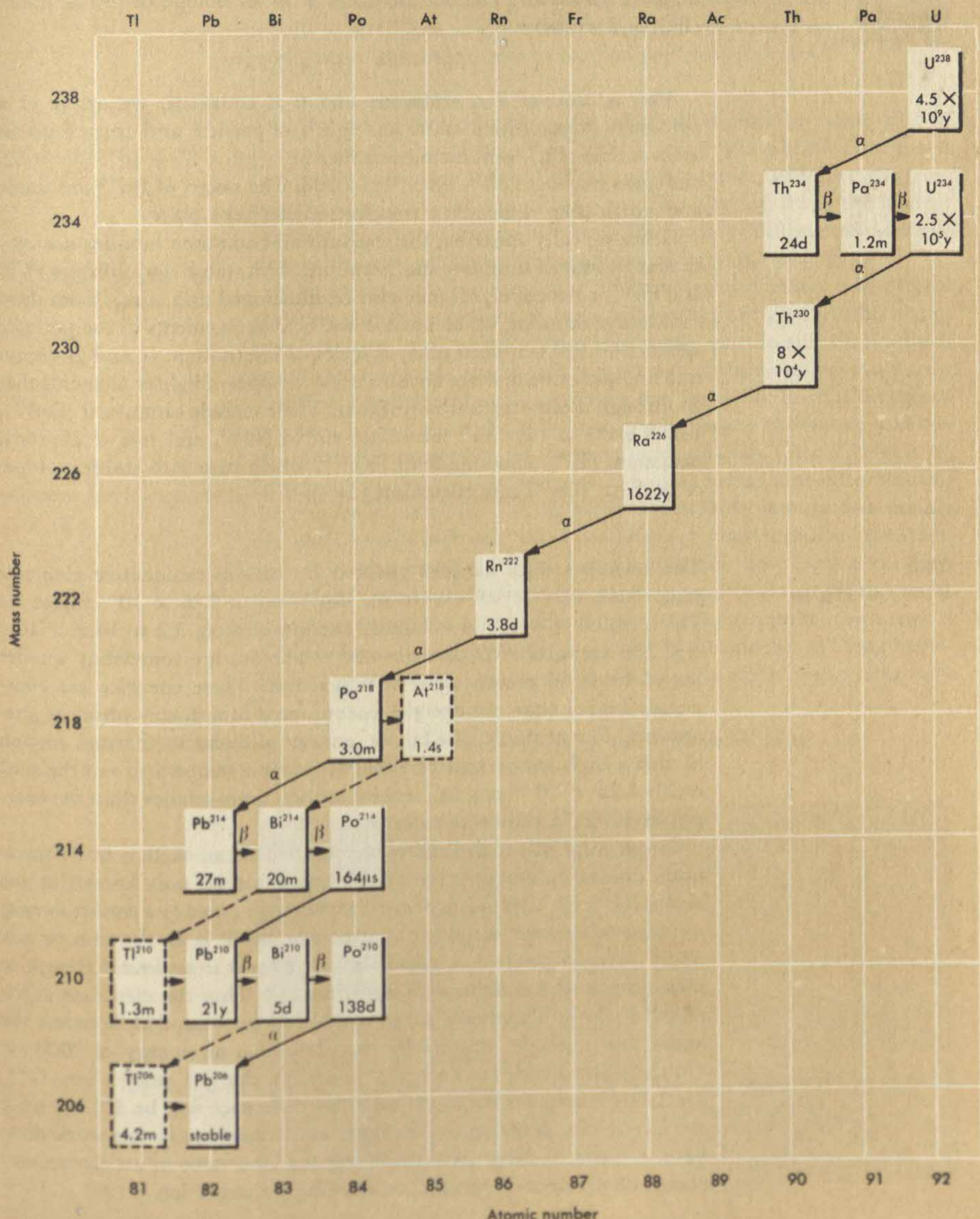
In the principal uranium family, which also includes radium, uranium plays the role of the head of the family, and being very long-lived, produces numerous children, grandchildren, great-grandchildren, etc. The genealogy of this family is shown in Fig. 26-3. The nucleus of an atom of ${}_{92}U^{238}$ emits an α particle and is transformed into another element which we can temporarily call X. Since the α particle carried away 2 protons and 2 neutrons for a total of 4 nucleons, we see that X must have an atomic number of $92 - 2 = 90$ and a mass number of $238 - 4 = 234$. The periodic table shows that the element of atomic number 90 is called thorium, Th. This transformation of an isotope of uranium into an isotope of thorium can be simply recorded as a nuclear equation:

$${}_{92}U^{238} \longrightarrow {}_{2}He^{4} + {}_{90}Th^{234}.$$

Since, in all nuclear reactions, charge (i.e., atomic number) is conserved, and since the number of nucleons (mass number) is conserved, it is apparent that the superscripts and subscripts must separately add up to the same values on both sides of the arrow.

Now the ${}_{90}Th^{234}$ is also unstable and radioactive, and it emits a β particle to become something else we may call Y—not to be confused here with the symbol for the element yttrium! The emission of this electron does not change the number of nucleons, so Y must have a mass number of 234. But the emission of the negative electron is the same thing as adding a positive charge, so that the atomic number of Y must

FIG. 26–3 The uranium-238 family. The numbers in the squares give the half-lives in years (y), days (d), hours (h), or seconds (s). Notice that near the end of the decay chain there are several alternative paths shown by dashed lines, which are taken by some of the decaying atoms.



Tl
Pb
Bi
Po
At
Rn
Fr
Ra
Ac
Th
Pa
U
238
234
230
226
222
218
214
210
206
Mass number
81
82
83
84
85
86
87
88
89
90
91
92
Atomic number
U238 4.5 × 10^9 y
α
Th234 24d
β
Pa234 1.2m
β
U234 2.5 × 10^5 y
α
Th230 8 × 10^4 y
α
Ra226 1622y
α
Rn222 3.8d
α
Po218 3.0m
At218 1.4s
α
Pb214 27m
β
Bi214 20m
β
Po214 164μs
Tl210 1.3m
Pb210 21y
β
Bi210 5d
β
Po210 138d
α
Tl206 4.2m
Pb206 stable

be 91. This atomic number identifies Y as an isotope of protactinium (Pa), and we have

$$_{90}\mathrm{Th}^{234} \longrightarrow {}_{-1}\beta^0 + {}_{91}\mathrm{Pa}^{234}.$$

After a total of 7 α emissions and 6 β emissions, we arrive at a polonium atom, which emits an eighth α particle and turns into an atom of lead (Pb), with atomic number $92 - (8 \times 2) + (6 \times 1) = 82$, and mass number $238 - (8 \times 4) = 206$. The nuclei of Pb^{206} are stable and no further radioactive transformations take place.

Genealogically speaking, the thorium and actinium families are very similar to that of uranium and terminate with stable lead isotopes Pb^{208} and Pb^{207}, respectively. It may also be mentioned that, apart from these radioactive families, which include the heaviest elements of the periodic system and are transformed by a series of intermittent α and β decays into isotopes of lead, there are also a few isotopes of lighter elements that go through a one-step transformation. These include samarium (Sm^{148}), which emits α rays and turns into stable Nd^{144}, and two β emitters, potassium (K^{40}) and rubidium (Rb^{87}), which turn into stable isotopes of calcium (Ca^{40}) and strontium (Sr^{87}).

26-6 Decay Energies

The velocities of α particles emitted by various radioactive elements range from 0.98×10^9 cm/sec for Sm^{148} up to 2.06×10^9 cm/sec for Th^{223}, which correspond to kinetic energies of from 3.2 to 14.2×10^{-6} erg. The energies of β particles and γ photons are somewhat smaller but of the same general order of magnitude. These energies are enormously higher than the energies encountered in ordinary physical phenomena. For example, the kinetic energy of atoms in thermal motion at such a high temperature as 6000°K (surface temperature of the sun) is only 1.25×10^{-12} erg, i.e., several million times smaller than the energies involved in radioactive decay.

In speaking about the energies liberated in radioactive transformations, nuclear physicists customarily use a special unit known as the *electron volt* (ev). This unit is defined as *the energy gained by a particle carrying one elementary electric charge* (no matter whether it is an electron or any singly charged positive or negative ion) *when it is accelerated through an electric field with a potential difference of one volt.* Thus the electrons accelerated in J. J. Thomson's tube, with 5000 volts applied between the anode and cathode, acquire by this definition an energy of 5000 ev. On the other hand, the energy of a doubly charged oxygen ion, O^{++}, accelerated through the same potential difference will be 10^4 ev, since the electric force acting on the ions, and consequently the work done by it, is twice as large. Remembering that the value of an elementary charge on an electron, proton, or any singly charged ion is 1.60×10^{-19}

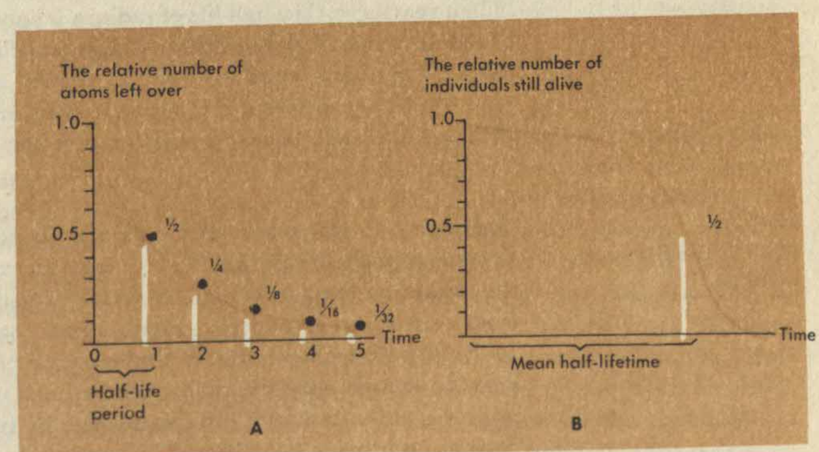
coul and that a volt is 1 joule/coul, *we find that one electron volt of energy is 1.60×10^{-19} joule, or 1.60×10^{-12} erg.* Also commonly used in nuclear work is the Mev, which stands for "million electron volts."

26-7 Half-lifetimes

The process of natural radioactive decay is ascribed to some kind of intrinsic instability of the atomic nuclei of certain chemical elements (especially those near the end of the periodic table), which results from time to time in a violent breakup and the ejection from the nucleus of either an α particle or an electron. The nuclei of different radioactive elements possess widely varying degrees of internal instability.

Whereas in some cases (such as U^{238}) radioactive atoms may remain perfectly stable for billions of years before they are likely to break up, in other cases (such as Po^{214}) they can hardly exist longer than a small fraction of a second. The breakup process of unstable nuclei is a purely statistical process, and we can speak of the "mean lifetime" of any given elements in just about the same sense as insurance companies speak of the mean-life expectancy of the human population. The difference is, however, that, whereas in the case of human beings and other animals the chance of decaying (i.e., dying) remains fairly low up to a certain age and becomes high only when the person grows old, radioactive atoms have the same chance of breaking up no matter how long it has been since they were formed. Since radioactive atoms begin to die out at the very moment of their birth, the decrease of their number with time is different from the corresponding decrease of the number of living individuals (Fig. 26-4A and B). In the latter case, the curve of surviving indi-

FIG. 26–4 A comparison of the survival curves for radioactive atoms and for living organisms.



viduals runs first almost horizontally and becomes steep later only when the organism begins to wear out; the radioactive decay curve is steep all the time.

The number of radioactive atoms that decay per unit of time is proportional to the number of atoms available, but is quite independent of the age of these atoms.

The time period during which the initial number is reduced to one-half is known as the "half-life" of the element. At the end of two half-lives, only a quarter of the original amount will be left; at the end of three half-life periods, only one-eighth will be left, etc. From the arguments above, we see a simple way to formulate mathematically the amount of a decaying element that is left after any number of half-lives. If we start out with an amount N_0 of some radioactive material, after n half-lives have passed, there will be left an amount N:

$$N = N_0 \times \left(\frac{1}{2}\right)^n.$$

The gas radon, for example, has a half-life of 3.8 days. If we start out with 5 mg of radon, how much will be left after a month? A month is $30/3.8 = 8$ half-lives, and

$$N = 5 \times \left(\frac{1}{2}\right)^8$$

$$= 5 \times \frac{1}{256} = 0.02 \text{ mg left.}$$

As we have seen above, various elements possess widely different lifetimes (Fig. 26-3). The half-life of U^{238} is 4.5 billion years, which accounts for its presence in nature in spite of the fact that all atoms of both stable and unstable elements may have been formed about 5 billion years ago. The half-life of radium is only 1620 yr, and the 200 mg of radium separated in 1898 by Marie and Pierre Curie now contains only 192 mg. The short-lived atoms of Po^{214} exist, on the average, for only 0.0001 sec between the moment they are formed by β decay of Bi^{214} and their subsequent transformation into Pb^{210} atoms.

26-8 Uranium-Lead Dating

The decay of radioactive elements and its complete independence of physical and chemical conditions gives us an extremely valuable method for estimating the ages of old geological formations. Suppose we pick up from a shelf in a geological museum a rock that is marked as belonging to the late Jurassic era, that is, to the period of the earth's history when gigantic lizards were the kings of the animal world. Geologists can tell approximately how long ago this era was by studying the thicknesses of various prehistoric deposits and by comparing them with the estimated

rates of the formation of sedimentary layers, but the data obtained by this method are rather inexact. A much more exact and reliable method, based on the study of the radioactive properties of igneous rocks, was proposed by Joly and Rutherford in 1913 and soon became universally accepted in historical geology. We have seen above that U^{238} is the father of all other radioactive elements belonging to its family and that the final product of all these disintegrations is a stable isotope of lead, Pb^{206}.

The igneous rock of the Jurassic era that now rests quietly on a museum shelf must have been formed as a result of some violent volcanic eruption of the past when molten material from the earth's interior was forced up through a crack in the solid crust and flowed down the volcanic slopes. The erupted molten material soon solidified into rock that did not change essentially for millions of years. But if that piece of rock had a small amount of uranium imbedded in it, as rocks often do, the uranium would decay steadily, and the lead resulting from that decay would be deposited at the same spot. The longer the time since the solidification of the rock, the larger would be the relative amount of the deposited lead with respect to the leftover uranium. Thus, by measuring the ratio of U^{238} to Pb^{206} in various igneous rocks, we can obtain very exact information concerning the time of their origin and the age of the geological deposits in which they were found.

Similar studies can be carried out by using the rubidium inclusions in old rocks and measuring the ratio of left-over rubidium to the deposited strontium. This method has an advantage over the uranium-lead method, because we deal here with a single transformation instead of the long sequence of transformations in the uranium family. In fact, one of the members of the uranium family is a gas (radon) and might partially diffuse away from its place of formation, thus leading to an underestimation of the age of the rocks.

26-9 Carbon Dating

Apart from the above-mentioned natural radioactive elements, which are presumably as old as the universe itself, we find on the earth a number of radioactive elements that are being continuously produced in the terrestrial atmosphere by cosmic-ray bombardment. Cosmic rays consist mostly of protons, plus α particles, and in smaller quantities the nuclei of more massive atoms. These high-energy particles, traveling at very nearly the speed of light, shower down constantly from all directions onto the earth's atmosphere. Their origin is not yet clear, but they may start from violent supernova explosions of distant unstable stars, to be further accelerated by passage through the weak but vast magnetic fields of space. Wherever they come from, their collisions with the atoms of our upper atmosphere give rise to the formation of showers of new

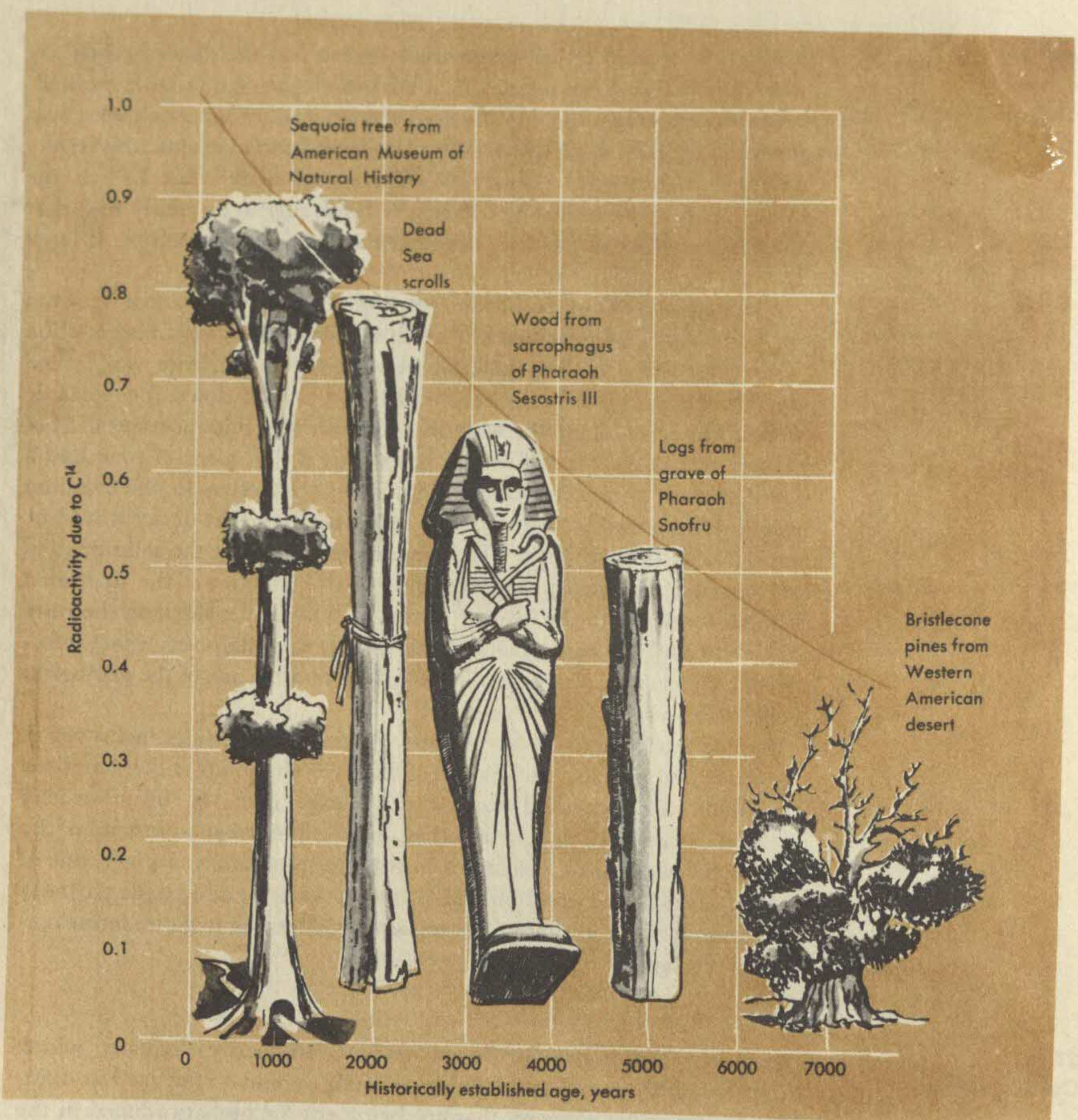


FIG. 26–5 The carbon-14 radioactivity of various old objects as a function of their age, as measured by Dr. Willard Libby.

particles, some from the shattered debris of the collisions and some created new from pure energy. Among these particles are many neutrons, which may themselves cause further changes. One of these changes is the reaction between a high-energy neutron and the nucleus of a common atom of atmospheric nitrogen (N^{14}). This produces an unstable isotope of carbon (C^{14}) and a proton:

$$_0n^1 + {}_7N^{14} \longrightarrow {}_1H^1 + {}_6C^{14}.$$

These constantly replenished C^{14} atoms are soon oxidized by the atmospheric oxygen to become incorporated into the molecules of atmospheric carbon dioxide. Since plants use atmospheric carbon dioxide for their growth, radioactive carbon is incorporated into each plant's body, making all plants slightly radioactive throughout their lives.

As soon as a tree is cut or falls down and all its metabolic processes stop, no new supply of C^{14} is taken in, and the amount of radioactive carbon in the wood gradually decreases as time goes on. Carbon-14 is a β emitter, and, as we can see, reverts back to ordinary nitrogen in the process:

$${}_6C^{14} \longrightarrow {}_{-1}\beta^0 + {}_7N^{14}.$$

Since the half-life of C^{14} is 5700 yr, the decay will last for many millennia; and by measuring the ratio of C^{14} to the ordinary C^{12} in old samples of wood, we are able to estimate dates of origin. The studies in this direction were originated by an American physicist, Willard Libby (1908–   ), and are playing the same role in the exact dating of ancient human history as the measurement of the uranium-lead ratio in the dating of the history of our globe. The measurement of C^{14} radioactivity in old samples of wood is a very delicate matter, since it is usually much weaker than the radioactivity of both the background surrounding the object (the experimenter himself has a higher C^{14} concentration than the piece of wood he is studying) and cosmic rays. Thus the sample under investigation must be heavily shielded, and a very sensitive counter must be used. In Fig. 26-5. we give a few examples of the measured and expected concentrations of radioactive carbon in various wooden objects of known age. Using these data and measuring the C^{14} concentration in wooden objects of unknown age, we can easily estimate their ages. Some examples of such estimates are given in Table 26-2.

Measurements of the C^{14} content of trees felled by ice-age glaciers have established that the last glaciation of the northern United States was much more recent (about 11,000 yr ago) than had been previously supposed.

26-10 Tritium Dating

Another method of dating by the use of radioactive materials, which was also worked out by Libby, utilizes the radioactivity of tritium, the heavy unstable isotope of hydrogen with atomic weight 3. Tritium is also produced in the terrestrial atmosphere by the action of cosmic radiation and is carried to the surface by rains. Tritium also decays by β emission, which converts it into ${}_2He^3$. However, tritium's half-life is only 12.5 yr, so that all age measurements involving this isotope can be carried out only for comparatively recent dates. It seems that the most interesting application of the tritium dating method is in the study of

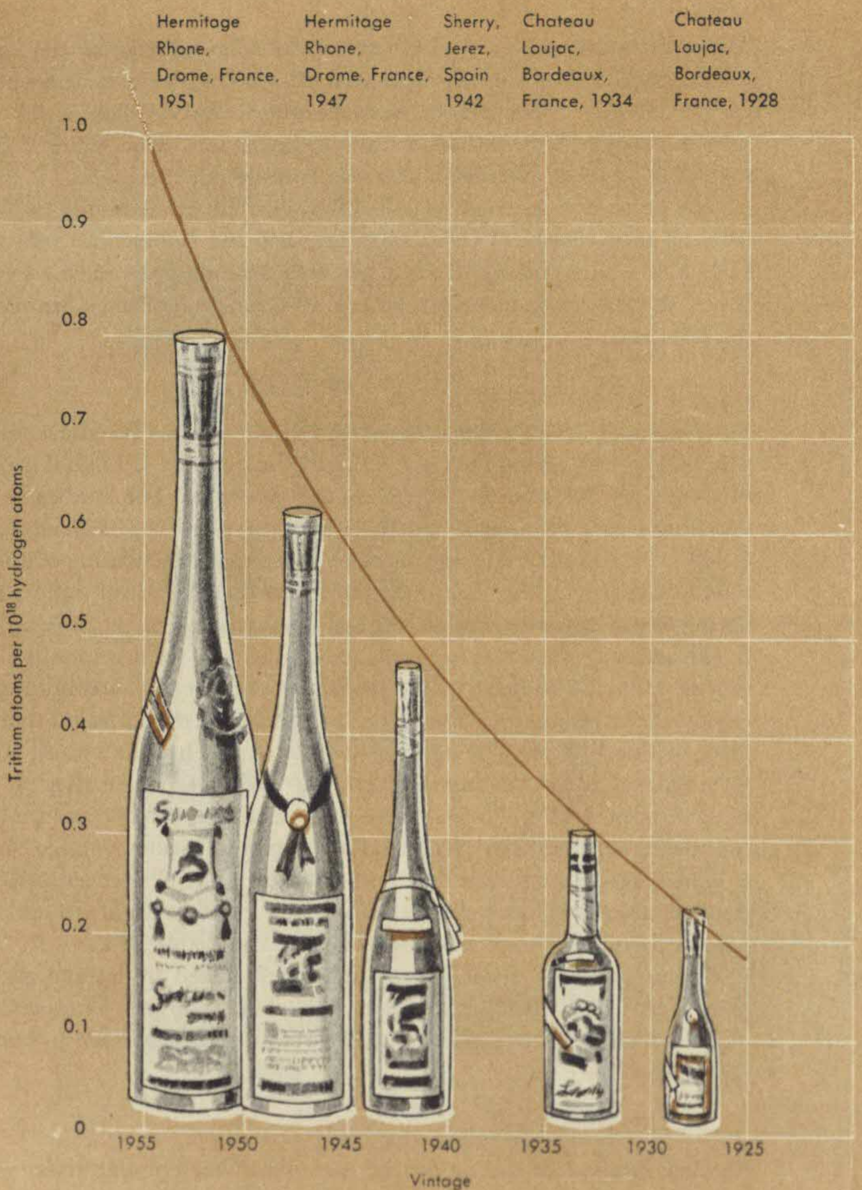


FIG. 26–6 The agreement between the calculated and the measured relationship between age and tritium activity in various wines.

the movements of water masses, both in ocean currents and in underground waters, since by taking samples of water from different locations and from different depths, we can tell by their tritium content how long ago this water came down in the form of rain.

Samples of old water are more difficult to collect than samples of old wood, and Libby resolved this problem by analyzing the tritium content in wine of different vintages, originating in different countries. The unpleasant part of this task is that an entire case of fine wine has to be used for each measurement and is rendered undrinkable in the process. But the agreement with the expected tritium content was in all cases excellent, as demonstrated in Fig. 26-6.

Questions

(26-2)

1. How many protons and how many neutrons are there in the nuclei of the following elements: carbon, bismuth, scandium, gold?

2. Give the numbers of neutrons and protons in the nuclei of: calcium, aluminum, radium, selenium.

(26-3)

3. What is the maximum possible velocity for a singly ionized neon ion that is part of a canal ray from a tube operating at 500 volts?

4. Singly ionized ions of sodium form a canal ray originating in a 300-volt tube. What is their maximum velocity?

5. From data in Table 26-1, calculate the atomic weight of chlorine.

6. From the data of Table 26-1, calculate the atomic weight of zinc.

7. Write, in the standard notation, the isotopes of oxygen and zinc.

8. List, using standard notation, the isotopes of nitrogen and chlorine.

(26-4)

9. The following isotopes are known to be unstable, and emit alpha particles. Give the designation (atomic number, chemical symbol, and mass number) of the isotopes into which they are transformed: (a) $_{89}Ac^{225}$, (b) $_{90}Th^{223}$, (c) $_{92}U^{238}$.

10. The following unstable isotopes decay by emitting an alpha particle. Write the proper designation of the isotopes into which this decay transforms them: (a) $_{94}Pu^{236}$, (b) $_{91}Pa^{226}$, (c) $_{88}Ra^{221}$.

11. The following unstable isotopes decay by beta emission. Designate the isotopes into which they are transformed: (a) $_{88}Ra^{225}$, (b) $_{86}Rn^{223}$, (c) $_{94}Pu^{243}$.

12. The isotopes listed below decay by emitting beta particles. Into what isotopes does this decay transform them? (a) $_{82}Pb^{211}$, (b) $_{83}Bi^{210}$, (c) $_{94}Pu^{243}$.

(26-5)

13. In the uranium decay series, radium ($_{88}Ra^{226}$) is formed, which emits an alpha particle to become radon (Rn). Write a nuclear equation for the transformation, showing the atomic number and mass number for this isotope of radon.

14. Along the chain of decays starting with $_{90}Th^{232}$, $_{84}Po^{216}$ is formed, which emits an alpha particle to become lead. Write a nuclear equation for this decay.

15. Potassium-40 often decays by β emission. Write the nuclear equation for this transformation.

16. An unstable isotope of chlorine, Cl^{36}, often decays by β emission. Write a nuclear equation describing this transformation.

17. Potassium-40 sometimes undergoes a transformation by another process than β decay. Instead of emitting an electron, the nucleus absorbs, or captures, one of the atomic electrons in the atom's K shell. Write a nuclear equation for this process, which is known as "K capture."

18. Chlorine-36 sometimes decays by K capture (see Question 17). Write the nuclear equation for this transformation.

19. Similar to the uranium family is that of thorium, which starts with ${}_{90}Th^{232}$ and ends with stable ${}_{82}Pb^{208}$. (a) How many alpha emissions are in this decay series? (b) How many beta emissions?

20. Another decay series does not exist in nature, but begins with the artificially produced ${}_{93}Np^{237}$. The series contains 8 α decays and 4 β decays. What is the designation of the stable isotope at the end of this series?

(26-6) 21. In the decay of radium into radon by alpha emission, gamma rays with wavelengths of 6.52×10^{-10} cm are also emitted. (a) What is this wavelength in Ångstrom units? (b) What is the energy of these gamma-ray photons, in Mev?

22. In the β decay of Pa^{234}, gamma rays of 0.043 Mev energy are also emitted. (a) What is the energy of these photons in ergs? (b) What is their wavelength, in cm and in Ångstroms?

23. What is the energy (in ev, ergs, and joules) of an electron that has been accelerated through a potential difference of 5×10^5 volts?

24. What is the energy in ev, ergs and joules, of an alpha particle that has been accelerated through a potential difference of 75,000 volts?

25. Take an electron and a doubly ionized nitrogen atom (N^{++}), and accelerate each from rest through a potential difference of 5000 volts. How will their kinetic energies compare?

26. A beam of alpha particles and a beam of ions of singly ionized neon are each accelerated through a potential difference of 3.1×10^4 volts. How do their kinetic energies compare?

(26-7) 27. Radon-222 has a half-life of 3.8 days. If we begin with 10.24 micrograms of this radon isotope, how much is left after 38 days?

28. Assuming that it were to remain undisturbed, calculate how much of Madame Curie's 200 mg of radium would be left in the year A.D. 8378.

29. An atomic bomb test produces 3.20 gm of a certain radioactive fission product that has a half-life of 4 months. How much time must elapse before the radioactivity of this particular product is reduced to less than 1 percent of its original intensity?

30. Strontium-90 (produced in appreciable quantities in atmospheric atom-bomb tests) has a half-life of 28 years. How long a time is needed for the initial Sr^{90} radioactivity to fall to 3 percent of its original intensity?

31. The parent of radium in the uranium series is ${}_{90}Th^{230}$, which has a half-life of about 8×10^4 yr. Suppose we are to set a gram of this pure thorium isotope aside for 160,000 years: (a) How much Th^{230} would be left? (b) Would we now have 0.75 gm of radium? (c) Would analysis show that some Pb^{206} was also present?

32. In the U^{238} decay series, Pa^{234} is produced by the β decay of Th^{234}, which has a half-life of 24 days. If 16 μgm of pure Th^{234} were put in a sealed container for 72 days: (a) How much Th^{234} would remain? (b) Would you expect to find 14 μgm of Pa^{234} in the container? (c) Would you expect there to be any Pb^{206} present?

(26-9) **33.** A piece of wood taken from an archaeological excavation has a ratio of C^{14} to C^{12} which is about one-fourth as large as the ratio in recently cut trees. How old is the piece of wood?

34. Charcoal from a long-buried ancient campfire has a C^{14}/C^{12} ratio about 12.5 percent as great as that found in modern wood. What is the approximate date of the ancient campfire?

(26-10) **35.** A diver inspecting a sunken ship finds in one of the cabins an unopened bottle of whiskey. The radioactivity of the whiskey due to tritium was found to be only about 3 percent of that measured in a recently purchased bottle marked "7 years old." How long ago did the ship sink? (The whiskey on the ship was moonshine, presumably made, bottled, and sunk in the same year.)

chapter / twenty-seven

Artificial Nuclear Transformations

27-1 Splitting Atomic Nuclei

After Rutherford became completely persuaded that the radioactive decay of heavy elements was due to the intrinsic instability of their atomic nuclei, his thoughts turned to the possibility of producing the artificial decay of lighter and normally stable nuclei by subjecting them to strong external forces. True enough, it was well known at that time that the rates of radioactive decay are not influenced at all by high temperatures or by chemical interactions, but this could be simply because the energies involved in thermal and chemical phenomena are much too small as compared with the energies involved in nuclear disintegration phenomena. Whereas the kinetic energy of thermal motion (at a few thousand degrees) as well as the chemical energy of molecular binding are of the order of magnitude of only 10^{-12} erg, the energies involved in radioactive decay are of the order of 10^{-6} erg, i.e., a million times higher. Thus, in order to have any hope of a positive outcome, the light stable nuclei must be subjected to a much stronger external agent than just a high temperature or a chemical reaction, and the bombardment of light

nuclei by high-energy particles ejected from unstable heavy nuclei was the natural solution of the problem.

Following this line of reasoning, Rutherford directed a beam of α particles emanating from a small piece of radium against nitrogen gas and observed, to his complete satisfaction, that besides the α particles that passed through the nitrogen and were partially scattered in all directions, there were also a few high-energy protons that were presumably produced in collisions between the onrushing α projectiles and the nuclei of nitrogen atoms. This conclusion was later supported by cloud chamber photographs, as we shall discuss in the next section. The capture of an α particle followed by the ejection of a proton increases the atomic number of the nucleus in question by one unit ($+2 - 1 = +1$) and its mass by three units ($+4 - 1 = +3$), transforming the original nitrogen atom ${}_7N^{14}$ into an atom ${}_8O^{17}$ of a heavier isotope of oxygen. We can express this reaction by the following nuclear equation:

$$ {}_7N^{14} + {}_2He^4 \longrightarrow {}_8O^{17} + {}_1H^1. $$

Following this original success, Rutherford was able to produce the artificial transformation of other light elements such as aluminum, but the yield of protons produced by α-bombardment rapidly decreased with increasing atomic number of the target material, owing to the increase in electrostatic repulsion of the α particle by the greater + charge of the larger nuclei, and he was not able to observe any ejected protons for elements heavier than argon (atomic number 18).

27-2 Photographing Nuclear Transformations

The study of nuclear transformations was facilitated by the ingenious invention of still another Cavendish physicist, C. T. R. Wilson. This device, known as the "Wilson chamber" or *cloud chamber*, permits us to obtain a photograph showing the tracks of individual nuclear projectiles heading for their targets and also the tracks of various fragments formed in the collision. It is based on the fact that whenever an electrically charged fast-moving particle passes through the air (or any other gas), it produces ionization along its track. If the air through which these particles pass is saturated with water vapor, the ions serve as the centers of condensation for tiny water droplets, and we see long, thin tracks of fog stretching along the particles' trajectories. The scheme of a cloud chamber is shown in Fig. 27-1. It consists of a metal cylinder C with a transparent glass top G, and a piston P, the upper surface of which is painted black. The air between the piston and the glass top is initially saturated with water (or alcohol) vapor, generally by a coating of moisture on top of the piston. The chamber is brightly illuminated by a light source S through a side window W. Suppose now that we have a small amount of radioactive material on the end of a needle N, which is placed near the thin window O.

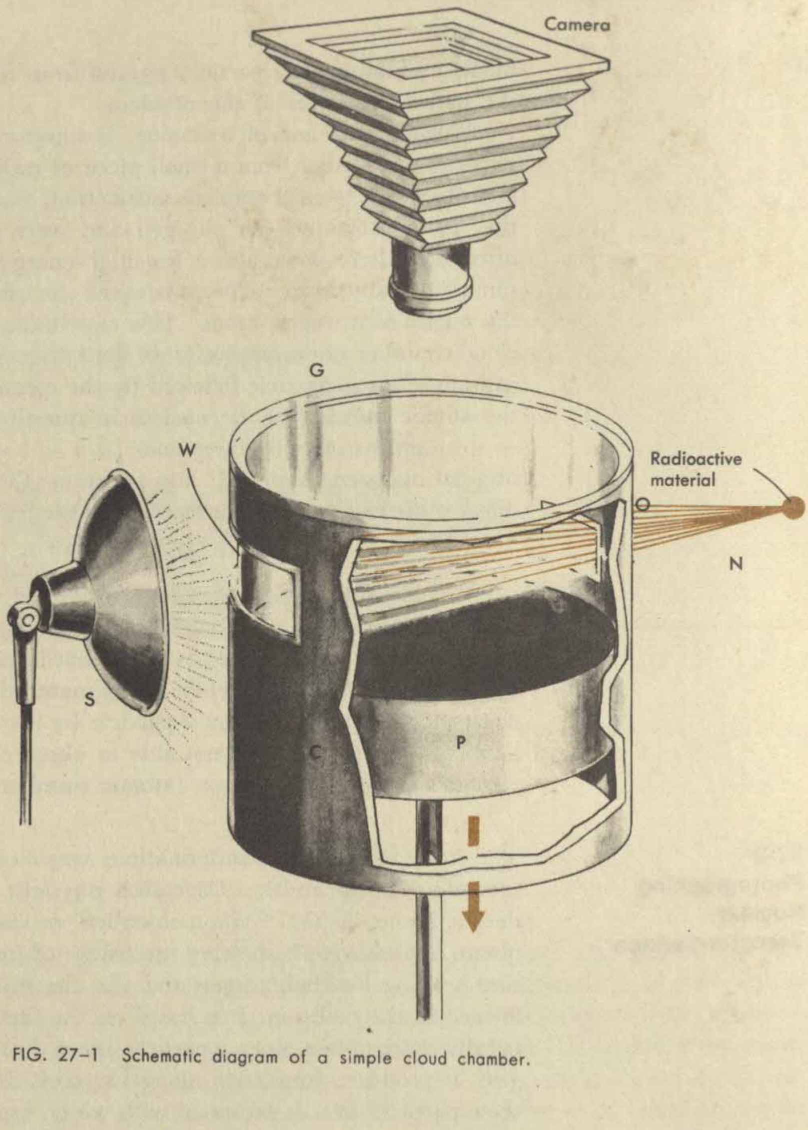
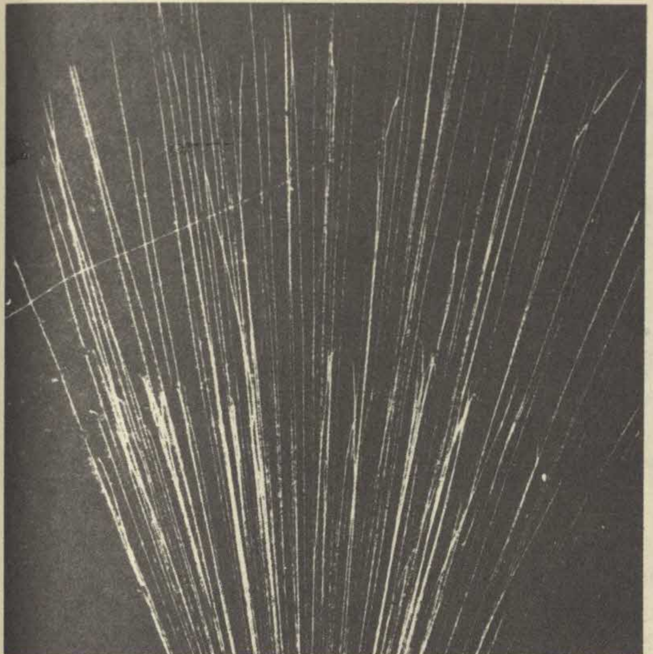


FIG. 27–1 Schematic diagram of a simple cloud chamber.

The particles ejected by the radioactive atoms will fly through the chamber, ionizing the air along their paths. However, the positive and negative ions produced by the passing particles recombine rapidly into neutral molecules. Suppose, however, that the piston is pulled rapidly down for a short distance. The rapid expansion of the air will cause it to cool, and the already saturated air now becomes supersaturated with moisture which will condense into water droplets. In order to condense, however, the droplets need centers of some sort around which to form. The natural condensation of raindrops takes place on dust particles, tiny salt crystals, or ice crystals. In *cloud seeding*, airplanes



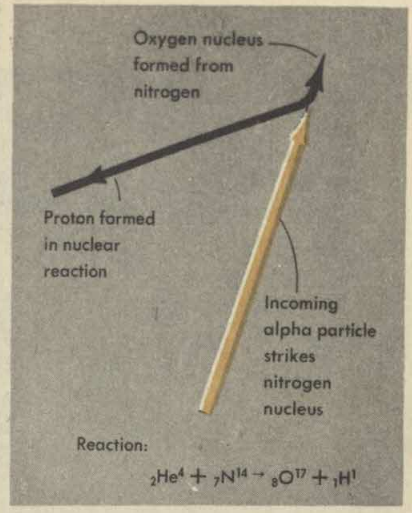


FIG. 27-2 The first cloud chamber photograph of a nuclear disintegration, taken by P. M. S. Blackett. The diagram to the right of the photo explains the tracks involved in this nuclear reaction.

scatter minute crystals of silver iodide to encourage the condensation of rain droplets.

In the cloud chamber, however, there is no dust, and the droplets condense on the ions (as the next best thing) that have been formed along the path of the speeding particles. Thus the tracks of the particles that passed by just before, or just as, the piston was pulled down will show up as trails of microscopic water droplets. The tracks of α particles are quite heavy, since the massive, doubly charged α particle ionizes the air strongly. The track of a proton is less strongly marked, and along the path of an electron, the ions and hence the droplets are much more sparse. In much of the present cloud chamber work, an intense magnetic field is created within the chamber, so the charged particles are deflected into curved paths. By measuring the curvature shown on the photographs, we can compute the speed and energy of the particles.

Figure 27-2 is a classical cloud chamber photograph taken in 1925 by P. M. S. Blackett that shows the collision of an incident α particle with the nucleus of a nitrogen atom in the air which fills the chamber. The long, thin track going almost backward is that of a proton ejected in that collision. It can be easily recognized as a proton track because protons are four times lighter than α particles and carry only half as much electric charge; therefore they produce fewer ions per unit length of their path than α particles. The short, heavy track belongs to the nucleus $_8O^{17}$ formed in the process of collision.

27-3 Bubble Chambers

In recent years, the *bubble chamber* has been developed to supplement the work of the cloud chamber. Although the general principle of its operation is the same as that of the cloud chamber, the bubble chamber filled with a liquid (often propane or, more recently, liquid hydrogen) which is kept exactly at its boiling-point temperature. A slight expansion will reduce the pressure on the liquid, and bubbles of vapor will form on the ions which have been produced in the liquid by passing particles. The bubble chamber has a great advantage when the collision events to be observed are relatively rare. In the closely packed atoms of a liquid, many more nuclear collisions will occur than in a gas, and the observer will stand a much better chance of photographing what he is looking for than he would with a cloud chamber. Figure 30-3 shows a photograph of tracks of particles in a bubble chamber.

27-4 First Particle Accelerators

Since the only massive projectiles emitted by the nuclei of natural radioactive elements are α particles, i.e., the nuclei of helium, it was desirable to develop a method for the artificial production of beams formed by other atomic projectiles, particularly beams of high-energy protons. According to theoretical considerations, the ease with which a bombarding particle penetrates into the structure of a bombarded atomic nucleus depends on the atomic number (i.e., the nuclear electric charge) of the element in question. The larger the atomic number, the stronger is the electric repulsive force opposing the approach of α particles to the nucleus and, consequently, the smaller are the chances of a demolishing collision. Since protons carry only half of the electric charge carried by an α particle, they were expected to be much better as atomic

FIG. 27-3 Sir John Cockcroft (left) and author Gamow discuss the possibility of breaking up the nucleus by using artificially accelerated protons. (Cambridge, England, 1929.)

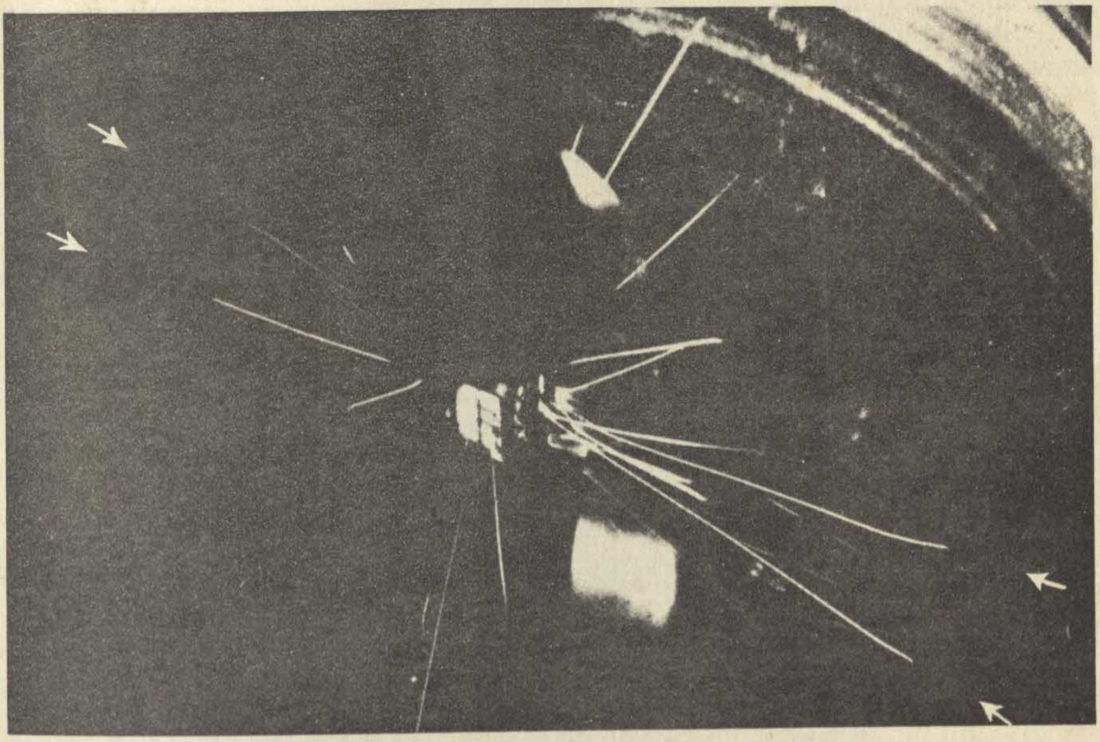


FIG. 27–4 A cloud chamber photograph of the breaking up of a lithium nucleus under the impact of an artificially accelerated proton (by P. Dee and E. T. S. Walton). The two pairs of diagonally opposite tracks, indicated by arrows, correspond to two pairs of alpha particles that resulted from the disintegration of two separate lithium nuclei.

projectiles and to be able to smash atomic nuclei of light elements even when moving with only 1 Mev (million electron-volts) of energy. Rutherford asked John Cockcroft (Fig. 27-3, left) to construct a machine that would accelerate protons to an energy of a million electron-volts, and, within a couple of years, the first "atom smasher" was constructed by Cockcroft and his associate E. T. S. Walton. Directing the beam of 1-Mev protons at a lithium target, Cockcroft and Walton observed the first nuclear transformation caused by artificially accelerated projectiles. The equation of this reaction is

$${}_3Li^7 + {}_1H^1 \longrightarrow 2\,{}_2He^4.$$

The two α particles resulting from the collision between the onrushing proton and the target, a lithium nucleus, are clearly seen in Fig. 27-4. If we use boron instead of lithium as the target, the reaction will be

$${}_5B^{11} + {}_1H^1 \longrightarrow 3\,{}_2He^4$$

FIG. 27-5 A cloud chamber photograph (by P. Dee and C. Gilbert), showing the three alpha particles (arrows) resulting from a boron nucleus struck by an artificially accelerated proton.

and the three α particles formed in this collision fly apart as shown in Fig. 27-5.

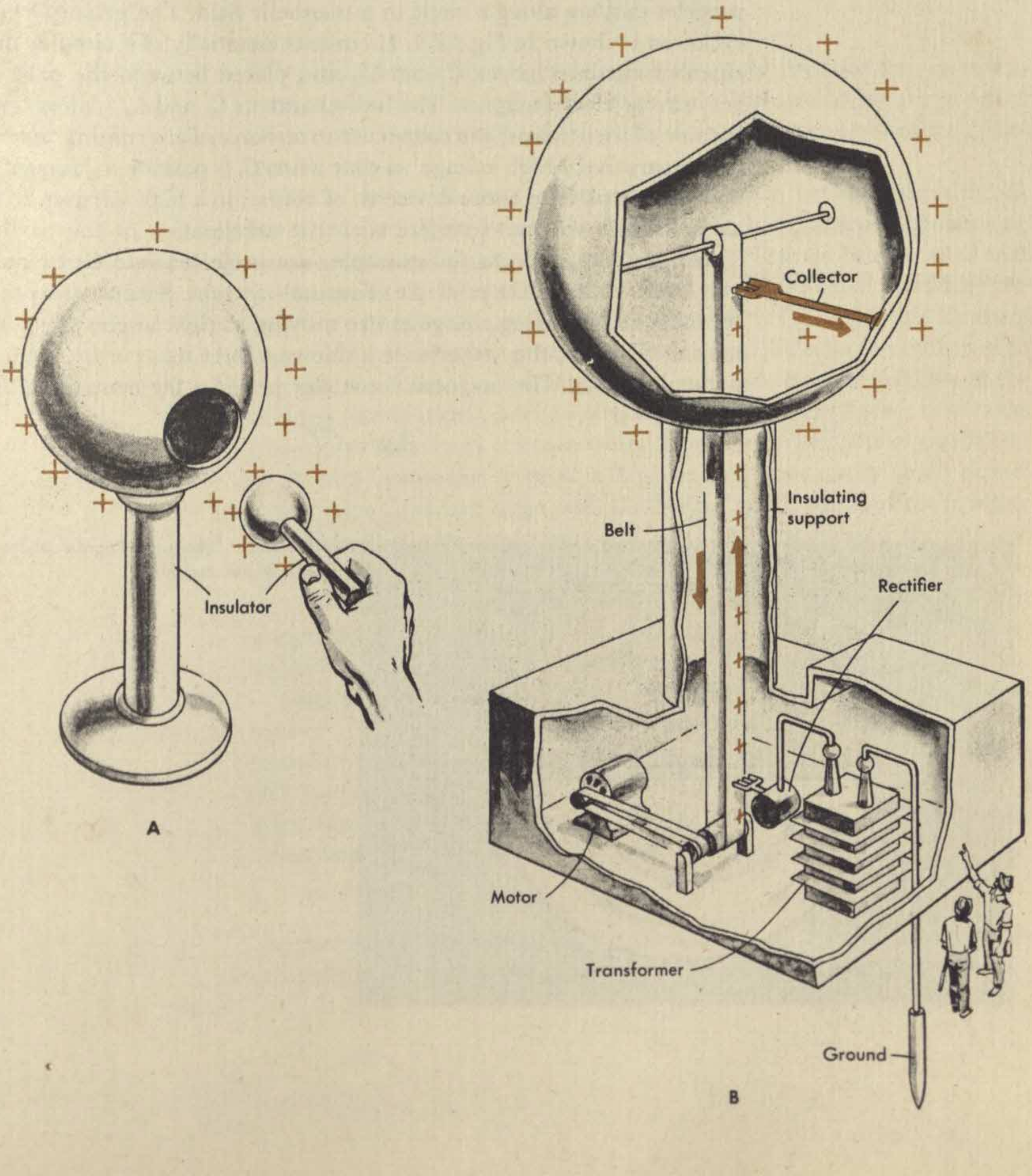
27-5 The "Van De Graaff"

Cockcroft and Walton's accelerator, which was based on the electric transformer principle, gave rise to a series of other ingenious devices for producing high-energy beams of atomic projectiles. The *electrostatic accelerator* constructed by R. Van de Graaff (1901–1967) and usually called by his name, is based on a classical principle of electrostatics, according to which an electric charge given to a spherical conductor is distributed entirely on its outer surface. Thus, if we take a hollow spherical conductor with a small hole in its surface, insert through this hole a small charged conductor attached to a glass stick, and touch the inside surface of the sphere (Fig. 27-6A), the charge will spread out to the surface of the big sphere.

Repeating the operation many times, we would be able to transfer to the large conductor any amount of electric charge, and raise its potential as high as desired (or, at least, until the sparks start jumping between the conductor and the surrounding walls).

In the Van de Graaff machine (Fig. 27-6B), the small charged ball is replaced by a continuously running belt that collects electric charges

FIG. 27–6 The basic principle (A), and the application of this principle in Van de Graaff's high-voltage electrostatic generator (B).



from a source at the base and deposits them on the interior surface of the large metallic sphere. The high electric potential developed in this process is applied to one end of an accelerating tube in which the ions of different elements are speeded up to energies of many millions of electron-volts.

27-6 The Cyclotron

Another popular atom smasher, called the *cyclotron*, was invented by E. O. Lawrence (1901–1958) and is based on an entirely different principle which utilizes the effect of many small accelerations of charged particles moving along a circle in a magnetic field. The principle of the cyclotron is shown in Fig. 27-7. It consists essentially of a circular metal chamber cut into halves, C_1 and C_2, and placed between the poles of a very strong electromagnet. The half-chambers C_1 and C_2 (called "dees" because of their shape) are connected to a source of alternating potential of comparatively high voltage, so that when C_1 is positive, C_2 is negative, and vice versa. The entire device is, of course, in a high vacuum so that air molecules will not interfere with the acceleration of the particles. These particles, protons for example, are projected into C_1 by an ion gun I that works on the principle of a canal-ray tube. Because the protons (or other ions bearing charge q) are moving at right angles to a strong magnetic field B, they experience a sideways force that curves them into a circular path. The magnetic force Bqv provides the centripetal force

$$Bqv = \frac{mv^2}{r}.$$

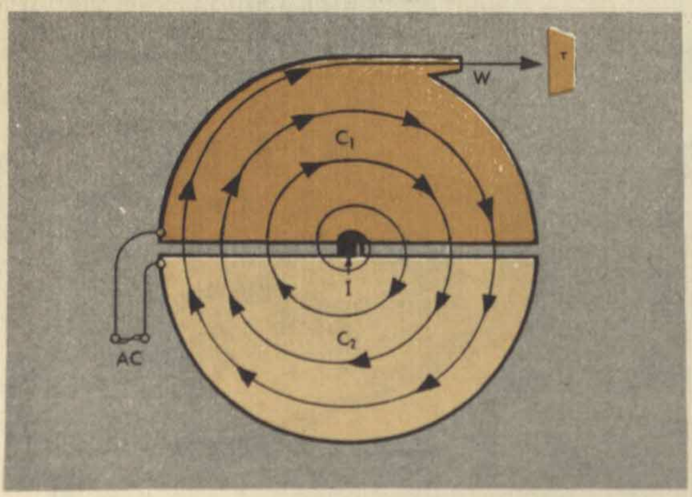


FIG. 27–7 Schematic diagram of the principle of the cyclotron.

The time required for the proton to complete a semicircle within C_1 is the length of the path πr divided by the velocity v:

$$t = \frac{\pi r}{v}.$$

From the centripetal force equation above, we can determine that

$$v = \frac{Bqr}{m}.$$

This value for v, substituted into the expression for t, gives us

$$t = \frac{\pi rm}{Bqr} = \frac{\pi m}{Bq}.$$

It is perhaps surprising to find that the time required for the ion to traverse a semicircle in one of the dees is not influenced by the speed of the ion nor by the radius of the path. This gives the secret of the cyclotron's success.

When the ion leaves C_1 and enters the space between the dees, C_1 has been made positive and C_2 negative, so that the space contains an electric field accelerating the ion toward C_2. (We must bear in mind that on the inside of either of the dees, there can be no electric field, so the only force here is that produced by the magnetic field.) Within C_2, then, the ion moves faster but describes a larger semicircle, so the time needed remains the same. On leaving C_2, the ion finds that the polarity of the dees has automatically been reversed, and so on. Every time it crosses the gap, a properly timed electronic oscillator has reversed the electric field, so the ion eventually receives a high energy from many small pushes rather than from one large one. Gathering speed, the ions move along an unwinding spiral trajectory and will finally be ejected through the window W in the direction of the target T. The field may oscillate millions of times per second, and at every oscillation a fresh squirt of ions is injected at I. In this way, the ions strike the target in little bunches millions of times a second, which almost amounts to a steady stream.

The largest existing cyclotron is in the radiation laboratory of the University of California. It has an accelerating circular box 60 inches in diameter and produces artificial α beams with an energy of 40 Mev (4.5 times higher than that of the most energetic natural α particles). With this atom smasher, it was possible to cause the artificial transformation of all elements up to the heaviest ones.

27-7 Other Particle Accelerators

Another look at the expression for the time a proton or other ion spends traversing a semicircle will reveal one drawback of the simple cyclotron:

$$t = \frac{\pi m}{Bq}.$$

The time (which is connected with oscillator frequency by the relation $f = 1/2t$) is seen to depend on m, the mass of the accelerated particle. For bullets or other particles traveling at ordinary speeds, this would not concern us, since m is practically constant. We know, however, from Einstein's Special Theory of Relativity that mass increases as particle speeds begin to approach the speed of light. So, if we try to speed up particles too much in a cyclotron, they will, in the outer spirals, be going so fast that their mass increase will throw them out of step with the oscillating field.

To eliminate this problem and permit particles to be accelerated to higher energies, many ingenious variations have been proposed and built: the synchrotron, the synchrocyclotron, the betatron, the proton synchrotron, and others. In general, these machines use changing oscillation frequencies, or changing magnetic field strengths, or both. We will not try to look into the details of their operation. Particles may be accelerated to nearly the speed of light and to energies measured in billions of electron volts.

When nuclear scientists try for higher and higher energies with these circular-path accelerators, they encounter another source of trouble. The necessary centripetal acceleration of the circling ions causes them (in accordance with Maxwell's equations) to radiate some of their energy as electromagnetic radiation. As added energy is given the particles, their increased centripetal acceleration causes most of it to be radiated away, thus defeating the purpose of the machine. There is one way to avoid this: centripetal acceleration $a_c = v^2/r$. Thus, for some given energy, the acceleration may be made less by making the radius of the path larger. This has led to the construction of some enormous accelerators, which have the handicap that cost increases rapidly with increasing size.

These difficulties have given a renewed interest to the linear, or straight-line accelerator. No huge magnets are needed when the particle path is straight. Stanford University now has a linear accelerator 2 miles long, in which electrons are accelerated to energies of 20 Bev (billion electron-volts). An electromagnetic wave is sent down a huge conducting tube with a speed only a trifle less than the speed of light. The electrons are pushed along by the traveling field of this wave like surfers riding the moving water wall of an ocean wave.

Questions

(27-1) **1.** In a nuclear reaction equation, an electron can be designated as $_{-1}e^0$ or $_{-1}\beta^0$, to indicate that its charge is -1, and that its absorption or emission does not change the mass number. Write the corresponding notation for a proton.

2. Write the designation of (a) a neutron, (b) an alpha particle, in a nuclear equation. (See Question 1.)

3. Write a nuclear equation for the following reaction: beryllium-9 is bombarded with alpha particles, and carbon-12 and neutrons are produced.

4. Write the following in the form of a nuclear equation: when an alpha particle strikes a boron-10 nucleus, a nitrogen isotope is produced, with the emission of a neutron.

5. Write nuclear equations for the following reactions: (a) Ordinary sodium, when bombarded with a proton, emits an α particle and becomes something else. (b) Tritium decays into something else by emitting a β particle.

6. Write the following nuclear reactions in the form of equations: (a) A certain isotope, under proton bombardment, emits an α particle and becomes ordinary sulfur. (b) Lithium-6, struck by a neutron, becomes an α particle and something else.

7. Complete the following reactions by giving the complete designation of the missing particle: (a) $B^{11} + H^1 \longrightarrow 3(He^4) + ?$ (b) $Al^{27} + H^2 \longrightarrow Mg^{25} + ?$ (c) $N^{14} + H^2 \longrightarrow 4\ (?)$. (d) $Al^{27} + H^2 \longrightarrow Al^{28} + ?$ (e) $Al^{27} + H^1 \longrightarrow \gamma + ?$

8. Complete the following reactions by giving the complete designation of the missing particle: (a) $Al^{28} \longrightarrow e^- + ?$ (b) $Al^{27} + He^4 \longrightarrow P^{30} + ?$ (c) $Al^{27} + H^2 \longrightarrow Si^{28} + ?$ (d) $P^{30} \longrightarrow e^+ + ?$ (e) $Al^{27} + n \longrightarrow Na^{24} + ?$

(27-4) **9.** What energy, in ergs and in joules, is acquired by a proton accelerated through a potential difference of 1 Mev?

10. An alpha particle is accelerated through a potential difference of 1 Mev. What is its energy in ergs and in joules?

(27-6) **11.** Deuterons are injected into a cyclotron whose magnetic field intensity is 8000 gauss (0.8 weber/m^2). The alternating potential difference between the dees is 4×10^4 volts. How many *revolutions* do the deuterons make before achieving an energy of 3 Mev? ($m_D = 3.34 \times 10^{-24}$ gm.)

12. A cyclotron has a magnetic field intensity of 6000 gauss (0.6 weber/m^2), and has an alternating potential difference of 5×10^4 volts between its dees. Protons are injected. How many *revolutions* must the protons make before obtaining a final energy of 4 Mev? ($m_P = 1.67 \times 10^{-24}$ gm.)

13. How long does it take the deuteron of Question 11 to traverse a semicircle within one of the dees?

14. How long does it take one of the protons of Question 12 to complete a semicircular trip through one of the dees?

15. What is the frequency of the alternating potential applied to the cyclotron dees of Question 11?

16. What is the frequency of the alternating potential applied to the dees of the cyclotron in Question 12?

17. What is the speed of a 3 Mev deuteron?

18. What is the speed of a 4 Mev proton?

19. What is the required diameter of the dees and magnet poles in the cyclotron of Question 11?

20. What is the required diameter of the dees and magnet poles of the cyclotron of Question 12?

(27-7) 21. If a cyclotron had been mistakenly designed to accelerate particles to half the speed of light, by what fraction would the time required to pass through a dee be increased, as compared with the time required at nonrelativistic speed?

22. A particle has a rest mass of m_0. By what fraction is its mass increased if accelerated to a speed of $0.75c$? By what fraction would the period of its revolution in a cyclotron be increased?

23. (a) By using the ordinary nonrelativistic expression for kinetic energy, figure the speed of an electron that has an energy of 5 Mev. (b) Is this possible? (c) Does this mean that it is impossible to give electrons 5 Mev of kinetic energy? Explain.

chapter / twenty-eight

The Structure of the Atomic Nucleus

28-1 Nuclear Particles

We have mentioned the early ideas holding that the nucleus was made up of protons and electrons—ideas that were rejected in the 1920s when quantum mechanics began its development. Rutherford, in 1920, without any definite experimental evidence to support him, had suggested a massive, electrically neutral nuclear particle, and had even given it its name, the *neutron*.

Actual evidence was not forthcoming, however, until 1930, when a German physicist, W. Bothe, noticed that the bombardment of beryllium by α particles from polonium gave rise to a very peculiar radiation of high penetrating power. Bothe believed that this radiation, unaffected by magnetic fields, was composed of high-energy photons. But in 1932, James Chadwick, a colleague of Rutherford, proved that the radiation must be composed of neutral particles of about the same mass as the proton. Rutherford had already provided the name for them, and it could now be seen that Bothe's original experiments could be written as

$$_4Be^9 + {}_2He^4 \longrightarrow {}_6C^{12} + {}_0n^1.$$

With this information, it was easy to assign to each isotope of each element the appropriate number of protons and neutrons in its nucleus. Although, when they are actually in a nucleus, neutrons seem to last forever, the situation is quite different for a free neutron outside a nucleus. Here, a neutron has a half-life of only 12 minutes and decays into a proton and an electron:

$$_0n^1 \longrightarrow {}_1H^1 + {}_{-1}e^0.$$

In some ways, neutrons are ideal projectiles to use in exploring the properties of nuclei. Being without electric charge, they are not subject to repulsive Coulomb forces and can penetrate even the heaviest nuclei with ease.

We notice that, whereas in the case of light nuclei the number of neutrons is equal to the number of protons, the number of neutrons exceeds the number of protons for heavier elements; the number of neutrons is 20 percent larger in iron and 50 percent larger in uranium. The preponderance of neutrons over protons in the heavier nuclei is due to the fact that, being positively charged, the protons repel each other, and their relative number must be reduced in order to secure the stability of the nucleus.

28-2 Models of the Nucleus

As we have seen before, it is hopeless to expect to be able to describe and understand the small world of photons and atomic particles in terms applicable to large-size phenomena we can directly see and touch. But such models—the wave picture of light, the Bohr atom, etc.—can be helpful if we recognize that they are only analogies and may not represent any "real" truth.

It is clear that the forces which hold the nucleus in one piece cannot be of a purely electric nature, since half of the nuclear particles (neutrons) do not carry an electric charge, whereas the other half (protons) are all positively charged, thus repelling one another and contributing to nuclear disruption rather than to stability. In order to understand why the constitutent parts of the nucleus stick closely together, we must assume that there exist between them forces of some kind, attractive in nature, which act on uncharged neutrons as well as on positively charged protons. These *nuclear forces*, indeed, appear to make no distinction whatever between neutrons and protons.

In Chapter 20, we discussed the ideas of surface tension and surface energy, that cause a bit of mercury or a raindrop to form a nearly spherical droplet, in order to make its surface as small as possible. This was the result of the cohesive forces between the molecules of the fluid.

The assumption that the forces acting between the constituent particles of the nucleus are analogous to those acting between the molecules of any ordinary liquid leads to the ***droplet model*** of an atomic nucleus,

according to which different nuclei are considered as minute droplets of a universal *nuclear fluid.*

The first important consequence of the nuclear-droplet theory is the conclusion that the volumes of different atomic nuclei must be proportional to their masses if the density of the fluid always remains the same, regardless of the size of the droplet which it forms. This conclusion is completely confirmed by direct measurements of nuclear radii by refined scattering methods, basically similar to Bohr's α-scattering experiments through which he discovered the nucleus. Thus the radii of the nuclei of oxygen and lead turn out to be 3×10^{-13} cm and 7×10^{-13} cm, respectively. Their actual masses (16 and 206 atomic mass units, respectively) are 2.7×10^{-23} gm and 3.4×10^{-22} gm. From these figures, we can calculate the densities of their nuclei. Since $d =$ mass/volume $=$ mass/$(\frac{4}{3}\pi r^3)$,

$$d_O = \frac{3 \times 2.7 \times 10^{-23}}{4\pi(3 \times 10^{-13})^3} = 2.4 \times 10^{14} \text{ gm/cm}^3$$

and

$$d_{Pb} = \frac{3 \times 3.4 \times 10^{-22}}{4\pi(7 \times 10^{-13})^3} = 2.4 \times 10^{14} \text{ gm/cm}^3.$$

Thus for both nuclei we get the same density of 2.4×10^{14} gm/cm^3, which is a very high density indeed! If the nuclear fluid, which is dispersed through matter in the form of minute droplets surrounded by rarefied electronic envelopes, could be collected to form a continuous material, one cubic centimeter of it would have a mass of 240 million metric tons!

As we saw when examining the periodic system of the chemical elements, the regular repetition of the chemical properties of atoms is due to the formation of consecutive shells and subshells in the electronic envelopes of the atoms.

Similar periodic changes are also observed in the case of atomic nuclei, manifesting themselves in the behavior of nuclear binding energies, magnetic properties, the ability to participate in various nuclear reactions, etc. One example of these periodic irregularities is shown in Fig. 28-1, in which the ability of different nuclei to capture additional neutrons (i.e., *neutron capture cross section*) is plotted against mass number. We notice that the regular increase of this ability with increasing atomic weight is interrupted by sharp minima.

Detailed studies of these and other variabilities of nuclear properties led to the conclusion that they always occur when either the number of neutrons or the number of protons is one of the following numbers: 2, 8, 14, 20, 28, 50, 82, 126, which represent the number of particles at which nuclear shells are completed. These so-called "magic numbers" are analogous to the sequence of numbers—2, 10, 18, 36, 54, etc.

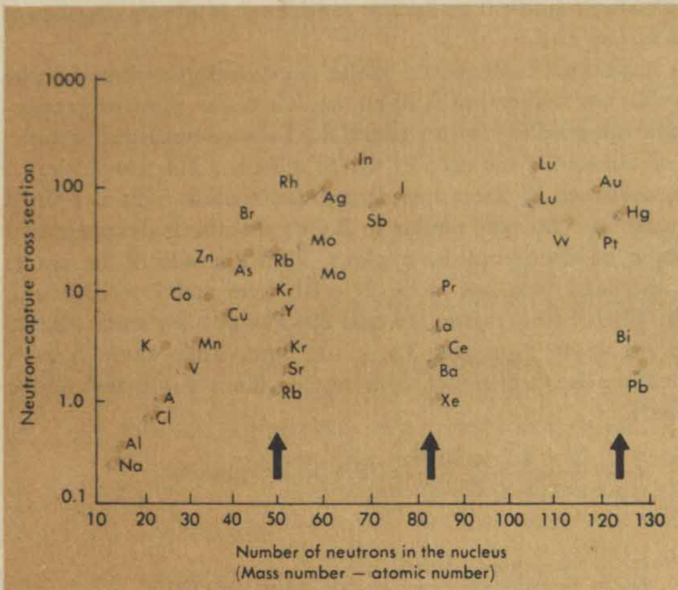


FIG. 28-1 Graph of neutron capture cross-section (a measure of the ability of nuclei to absorb colliding neutrons).

(atomic numbers of the rare gases)—that characterize the periodic system of chemical elements and that represent the number of electrons at which atomic shells are completed. The abnormally small neutron-capture cross sections for the elements with completed neutron shells (Fig. 28-1) are analogous to the chemical inertness of the rare gas atoms that possess completed electron shells.

There are, however, two important differences between the shell structure of nuclei and the shell structure of atoms. In atoms, one system of shells accommodates the electrons of the atomic envelope, while in nuclei there are two independent sets of shells: one for neutrons and one for protons. Another difference lies in the fact that, whereas the electron shells in the atom are geometrically separated, nuclear shells apparently interpenetrate each other and so can be distinguished only by their different energies.

28-3 Mass Defect and Nuclear Binding Energy

In comparing the masses of various atomic nuclei with the masses of the protons and neutrons from which they are formed, we always find a slight discrepancy. For example, in the case of the oxygen isotope O^{16}, we have

$$
\begin{aligned}
8 \text{ neutrons} &= 8 \times 1.00867 = 8.06936 \\
8 \text{ protons} &= 8 \times 1.00728 = 8.05824 \\
8 \text{ electrons} &= 8 \times 0.00055 = \underline{0.00440} \\
& \qquad\qquad\qquad\quad 16.13200
\end{aligned}
$$

which may be compared with the atomic weight of O^{16}, which is 15.99491.

It was necessary to add in the 8 electrons because atomic weights always include the weight of the atomic electrons in their values. Since the number of atomic electrons is always equal to the number of protons in the nucleus, it may be more convenient, in figuring nuclear discrepancies, to use the hydrogen atom instead of the proton, as this automatically includes the proper number of electron masses. For oxygen, we could have written

$$\begin{aligned} 8 \text{ neutrons} &= 8 \times 1.00867 = 8.06936 \\ 8 \text{ hydrogens} &= 8 \times 1.00783 = \underline{8.06264} \\ & \ 16.13200 \end{aligned}$$

Thus, by either method, the oxygen nucleus is seen to be lighter than its constituents by 0.13709 units of atomic weight, or atomic mass units (amu).

Similarly, in the case of the principal isotope of iron, we have

$$\begin{aligned} 30 \text{ neutrons} &= 30 \times 1.00867 = 30.26010 \\ 26 \text{ hydrogens} &= 26 \times 1.00783 = \underline{26.20358} \\ & \ 56.46368 \end{aligned}$$

which is to be compared with the value of 55.9349 for the atomic weight of Fe^{56}. Here we find that the atomic weight of Fe^{56} is 0.5288 amu smaller than the combined masses of its components.

The explanation of this *mass defect* lies in Einstein's $E = mc^2$, and might perhaps have better been called an "energy defect." Since the mass of, say, the O^{16} nucleus is 0.13709 amu less than the combined masses of the 8 neutrons and 8 protons that make it up, it must therefore have the equivalent of 0.13709 amu less energy than the 16 nucleons considered separately. It is easy to see (at least qualitatively) that this must be so. The separated nucleons have great potential energy because of the strong attractive nuclear forces between them. Or we may look at the process in reverse, and imagine ourselves picking the nucleus apart. Every nucleon that is removed from it will have to be dragged away against the strong attraction of its neighbors. When the 16 nucleons (for O^{16}) have been separated in this way, the work we had to do will be just equal to the added potential energy of the separated nucleons. This work, or energy, is the *binding energy* of the nucleus. For the oxygen nucleus, we see that this total binding energy is equivalent to a mass of 0.13709 amu.

This mass can readily be converted to other units. Since the amu, the unit of atomic mass ($\frac{1}{12}$ of the actual mass of the C^{12} isotope of carbon) equals 1.66×10^{-24} gm, its energy-unit equivalent is

$$1.66 \times 10^{-24} \times (3 \times 10^{10})^2 = 1.48 \times 10^{-3} \text{ erg},$$

and

$$\frac{1.48 \times 10^{-3}}{1.60 \times 10^{-12}} = 9.31 \times 10^{8} \text{ ev.}$$

We thus have the following convenient conversions:

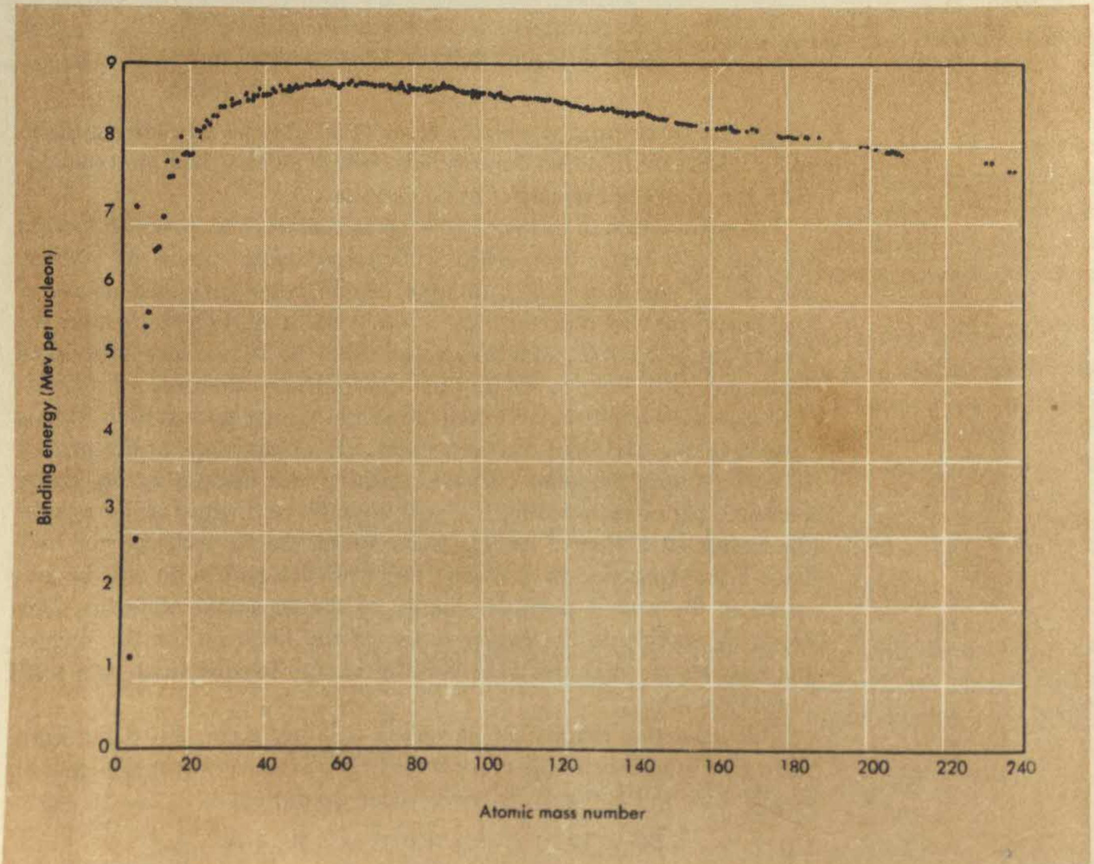
$$1 \text{ amu} = 1.48 \times 10^{-3} \text{ erg} = 1.48 \times 10^{-10} \text{ joule}$$

or

$$1 \text{ amu} = 9.31 \times 10^{8} \text{ ev} = 931 \text{ Mev.}$$

It is often useful to divide the total binding energy of a nucleus by the number of nucleons it contains, to get the *binding energy per nucleon.* Figure 28-2 shows a plot of the binding energy per nucleon, plotted against mass number for a representative sample of elements. This binding energy per nucleon (which has relatively small values for light

FIG. 28-2 The nuclear binding energies, per nucleon, of stable and long-lived nuclei.



nuclei) increases with mass number, reaches a maximum in the neighborhood of mass number 50, and then decreases again for the more massive nuclei.

28-4 Mass Defect and Nuclear Reactions

The exact knowledge of the atomic weights of the isotopes permits us to evaluate the energy balance of various nuclear reactions, since the mass equivalent of the liberated or absorbed nuclear energy must enter into the equation of the conservation of mass during the transformation. Thus, in the case of Rutherford's original reaction,

$$_7N^{14} + {_2He^4} \longrightarrow {_8O^{17}} + {_1H^1}.$$

The sums of the masses of the atoms entering the reaction and those resulting from it are, respectively,

$$\begin{array}{lr@{\qquad}lr} _7N^{14} & 14.00307 & _8O^{17} & 16.99914 \\ + & & + & \\ _2He^4 & 4.00260 & _1H^1 & 1.00783 \\ \hline & 18.00567 & & 18.00697 \end{array}$$

The combined mass of the reaction products is larger than the combined mass of the atoms entering into the reaction, by 0.00130 amu, indicating that this much energy has been converted into mass in the process. Using the mass-energy conversion factor given in a previous section, we obtain for the energy balance

$$\Delta E = -1.84 \times 10^{-6} \text{ erg} = -1.15 \text{ Mev}$$

which coincides with the difference between the kinetic energy of the incident α particle and the kinetic energy of the ejected proton as observed by Rutherford. On the other hand, Cockcroft and Walton's reaction,

$$_3Li^7 + {_1H^1} \longrightarrow 2{_2He^4}$$

leads to the following:

$$\begin{array}{lr@{\qquad}lr} _3Li^7 & 7.01601 & _2He^4 & 4.00260 \\ + & & + & \\ _1H^1 & 1.00783 & _2He^4 & 4.00260 \\ \hline & 8.02384 & & 8.00520 \end{array}$$

In this case the difference is 0.01864 atomic mass units, corresponding to an energy liberation of 2.8×10^{-5} erg or 17.7 Mev per reaction. This energy shows up as kinetic energy of the newly created pair of α particles, which can be readily measured by their tracks in a cloud chamber.

28-5 Fusion and Fission

In considering atomic nuclei of different chemical elements as minute droplets of universal nuclear fluid, we may expect that these nuclear droplets will behave in somewhat the same way as droplets of ordinary liquid. In observing droplets of, let us say, mercury rolling on the surface of a saucer, we notice that whenever two droplets meet they fuse together, forming a larger droplet. The fusion of two droplets into one is the work of surface-tension forces, which tend to reduce the total free surface and surface energy of the liquid. In fact, it is easy to show that the surface of one big droplet is smaller than the combined surfaces of two half-size droplets. Since the total volume of the compound droplet is twice the volume of each of the smaller ones, its radius must be $\sqrt[3]{2} = 1.26$ times larger, and its surface $(1.26)^2 = 1.59$ times larger. So when two half-size droplets fuse into one, the total surface is reduced in the ratio 2 : 1.59, or by 20 percent. Thus the fusion of two droplets into one always leads to the liberation of surface energy and always takes place spontaneously whenever two droplets come into contact. If the surface-tension forces were the only forces acting in atomic nuclei, any two nuclei would fuse together, liberating nuclear energy.

However, the situation changes quite considerably if we take into account that electric forces of repulsion are also present in the nuclei. In contrast to ordinary liquids, a nuclear fluid is always electrically charged, since about half of its constituent particles are protons. The electric repulsion between the nuclear charges acts in the opposite direction to the surface-tension forces and tends to disrupt larger droplets into smaller ones, i.e., to cause *fission*, rather than fusion.

So, with both kinds of forces present, the answer to the question whether nuclear energy will be liberated in fusion or fission depends on the relative strength of the two forces. If we proceed along the sequence of elements from the lighter nuclei to the heavier ones, the surface energy, which is determined by the total surface of the nucleus, increases comparatively slowly, being proportional to the two-thirds power of the atomic weight. On the other hand, electric energy increases approximately as the square of the nuclear charge.

For light nuclei, the surface-tension energies (which give a release of energy by fusion) overshadow the effects of electric charge, so *the fusion of two light nuclei will liberate excess energy* as a by-product. However, since electric energy increases with atomic weight much faster than the surface energy does, we should expect the situation to be reversed for heavy nuclei, so that the electric charge factor is of greater importance. For these nuclei of high atomic number, we would thus expect the energy that is released by splitting the electric charge of the nucleus in two to be greater than the energy used up because of the greater surface area of the two fragments. This reasoning would lead us to predict that *excess energy will be released when a large nucleus fissions, or splits in two.*

This theoretical conclusion is in complete agreement with the empirical evidence given by the study of nuclear binding energies. Looking at Fig. 28-2, we notice that the value of the binding energy per particle increases with the atomic weight for the elements in the earlier (lighter) part of the periodic system. This means that if two nuclei belonging to this region fuse together, a certain amount of nuclear energy will be set free. On the other hand, for the later (heavier) part of the periodic system, the binding energy per particle decreases with increasing atomic weight, indicating that for these elements fission and not fusion will be the energy-liberating process. Between these two regions lie the elements in the neighborhood of the iron group, which have the maximum binding energy per particle and are therefore stable with respect to both fusion and fission.

28-6 Nuclear Potential Barrier

When a positively charged nuclear projectile such as an α particle or a proton approaches an atomic nucleus, it is acted upon by electrostatic repulsive forces and cannot come into direct contact with the nuclear attractive forces unless its kinetic energy is large enough to overcome the repulsion. However, as soon as the contact is achieved, nuclear attractive forces take hold of the approaching particle and pull it into the nucleus. Plotting the potential energy of a positively charged particle in the neighborhood of the atomic nucleus, we obtain the curve shown in Fig. 28-3. This curve represents a 'potential barrier" for the penetration of the incident positively charged particles into the nucleus, as well as for the

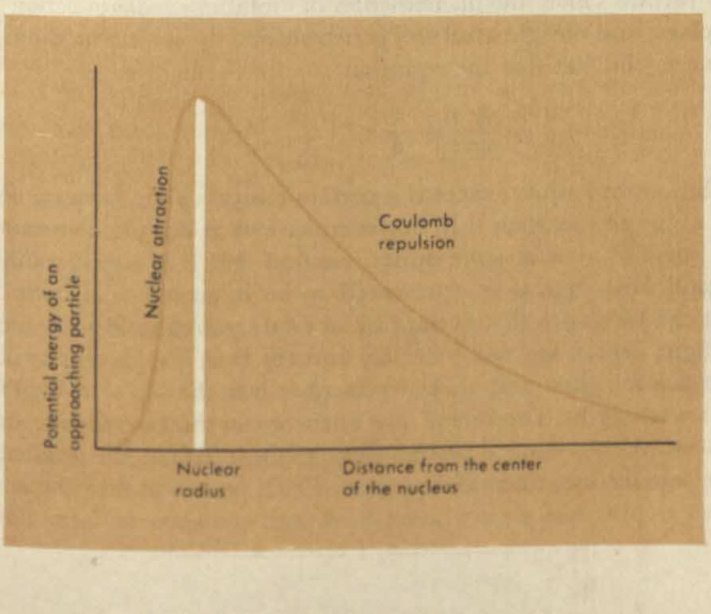


FIG. 28-3 The "potential barrier," represented by the height of this graph of the potential energy of a positively charged particle (alpha particle, proton, etc.) in the neighborhood of an atomic nucleus.

escape of such particles from the nucleus. According to classical mechanics, the incoming and outgoing nuclear particles can pass the potential barrier only if their kinetic energy is larger than the maximum height of the barrier. Experimental evidence shows, however, that this is definitely not so. The uranium nucleus, for example, has a radius of 9×10^{-13} cm and is surrounded by a potential barrier 27 Mev high. We would expect, therefore, that only particles having 27 Mev of potential energy or more would be able to escape from the uranium nucleus. We know, however, that α particles emitted by uranium have an energy of only 4 Mev, and it is difficult to understand how they get out across the barrier at all. Also, in the case of the artificial transformation of elements, such as Rutherford's experiments on the bombardment of nitrogen by α particles, the energy of the projectiles is often lower than the height of the potential barrier surrounding the bombarded nucleus; nevertheless, some of these projectiles penetrate into the nuclear interior, causing its disintegration.

28-7 Tunnel Effect

The paradoxical phenomenon just discussed, known as the *tunnel effect*, was explained in 1928 by G. Gamow and, independently, by R. Gurney and E. Condon, as being due to the wave nature of nuclear particles. In order to understand the situation, let us consider a simple example from the field of optics. As was discussed in Chapter 15, a light beam falling on the interface between a dense and a light material (passing from glass into air for example) will be refracted with an angle of refraction larger than the angle of incidence. If, however, the angle of incidence exceeds a certain value, the phenomenon of "total internal reflection" will take place, and no light at all will penetrate into the second medium. This is due to the fact that the equation

$$\frac{\sin i}{\sin r} = \frac{1}{n} < 1$$

has no solution for r when i exceeds a certain critical value, because no angle has a sine greater than 1. If, however, we look at this phenomenon from the point of view of wave optics, we find that it is considerably more complicated than it is represented to be in geometrical optics. Indeed, it can be shown that in the case of total internal reflection, the incident light waves are not reflected entirely from the geometrical surface separating glass and air but penetrate into the air to a depth of several wavelengths. The lines of flow of energy for this case are shown in Fig. 28-4A. We see that, on passing through the interface, the original light beam breaks into many components which penetrate into the air to different depths but always come back into the glass to form the

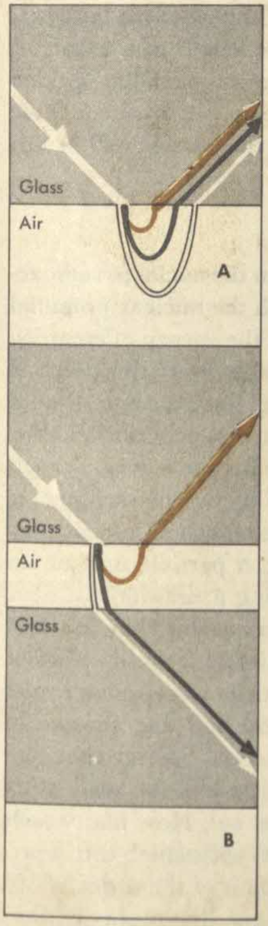


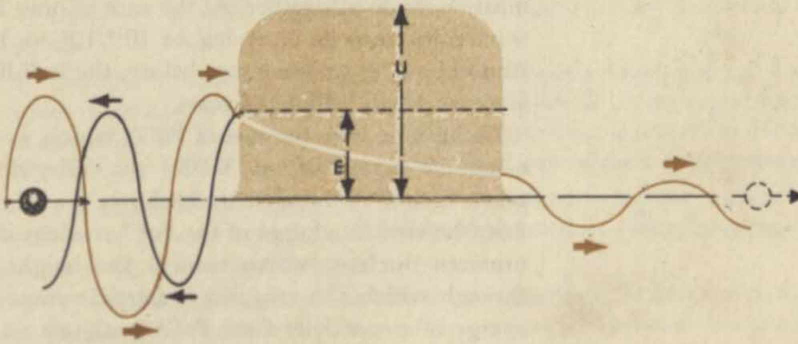
FIG. 28-4 The total internal reflection of light from a glass-air interface (A), and a partial penetration into another piece of glass placed within a few wavelengths of the first one (B).

reflected beam. This phenomenon cannot be described in terms of geometrical optics, and should be considered as a peculiar case of the diffraction of light.

If we place another piece of glass right under the first one (Fig. 28-4B) so that the thickness of the air layer between them will be equal to only a few wavelengths, some of the light entering the air layer will reach the surface of the second piece of glass and form in it a light beam parallel to the incident beam. The intensity of that beam decreases very rapidly with increasing thickness of the air layer, and becomes negligibly small when this thickness exceeds several wavelengths. Thus wave optics explains a phenomenon which would be completely unexplainable from the point of view of classical geometrical optics.

Let us now consider a material particle with an initial energy E, which falls on a potential barrier the height of which, U, is larger than the energy of the particle. According to classical mechanics, this particle cannot enter the region occupied by the potential barrier. The situation is different, however, if we consider the motion of the particle to be represented by de Broglie waves; the potential barrier plays the same role for the de Broglie waves as the air layer between the two pieces of glass plays for the light waves in the case of total internal reflection. The de Broglie waves incident on a potential barrier will be partially reflected from its outer boundary, but part of them will penetrate into the barrier itself (Fig. 28-5). The part that penetrates into the barrier will reach its other side and will come out into the region beyond the barrier. Since the propagation of de Broglie waves guides the motion of material particles, it follows that some of the particles falling on the barrier will pass through it even though this contradicts classical mechanics.

FIG. 28-5 The de Broglie wave of a particle falling on a potential barrier. The height of the barrier U is greater than E, the kinetic energy of the particle. The wave is partially reflected and partially passes through the barrier.



It should be noticed that the intensity of the de Broglie waves that pass through the barrier will become very small if the length of the barrier exceeds several wavelengths. Since the number of particles guided by de Broglie waves is proportional to their intensity, we must conclude that the number of particles which pass through the barrier will in this case also be very small.

28-8 Alpha Decay and Nuclear Bombardment

As we have mentioned, α particles can escape from the nuclei of radioactive elements only if they are able to pass through the nuclear potential barriers, the height of which exceeds many times the energy of escaping particles. Since such a feat is absolutely impossible from the point of view of classical mechanics, the phenomenon of radioactivity would not exist if the laws of classical mechanics were completely unshakable. We have seen, however, that the leakage of the de Broglie waves through potential barriers of any height opens the possibility for the escape of α particles from the nuclei, even though the chances of such an escape may be extremely low. It has been calculated that an α particle making an attempt to cross the potential barrier surrounding a uranium nucleus has only 1 chance in 10^{38} to do so. Incredibly small as this chance seems to be, success can be finally achieved if a sufficiently large number of attempts is made. Let us imagine that the α particles imprisoned in the nuclear interior are restlessly rushing to and fro and are constantly bouncing from the high walls of the nuclear potential barrier that surrounds them on all sides. Each time an α particle hits the wall of its prison, it has a slight chance (1 out of 10^{38}) to get out. How many such escape attempts are made per second? The velocity with which the imprisoned α particles move inside their nuclear prison is of the order of 10^9 cm/sec, while the size of the prison is about 10^{-12} cm. By simple division, we find that an α particle imprisoned within a nucleus collides with the surrounding walls about 10^{21} times per second. Since the chance of escape in any single collision is only 1 out of 10^{38}, 10^{38} escape attempts must be made altogether. At the rate of only 10^{21} attempts per second, α particles must go on trying for $10^{38}/10^{21} = 10^{17}$ sec $= 3 \times 10^9$ years. And indeed, as we have seen before, the half-lifetime of uranium nuclei is measured in billions of years.

Let us now take the case of Po^{214}, which, in contrast to uranium, has a half-life period of only 0.0001 sec. Why does this nucleus decay so much faster? A detailed study shows that there are two reasons for it: first, the electric charge of the Po^{214} nucleus is smaller than that of the uranium nucleus, which reduces the height of the potential barrier through which the escaping α particle must penetrate; secondly, the energy of α particles from Po^{214} is almost twice as large as that of α particles from uranium. Carrying out the same kind of calculations as

we did in the case of the uranium nucleus, we find that an α particle escaping from a Po^{214} nucleus must make only 10^{17} attempts to have a good chance to get away. At the rate of 10^{21} attempts per second, the mean waiting time for the escape reduces to $10^{17}/10^{21} = 0.0001$ sec, which agrees with the data given in the chart of Fig. 26-4.

Thus we see that comparatively small variations in the height of the potential barrier and in the velocity of escaping α particles can change the half-lifetime periods from billions of years to small fractions of a second. This accounts for the great variability of the half-life periods among the radioactive elements.

What is true for the α particles escaping from radioactive nuclei is also true for α particles and other positively charged atomic projectiles that are shot at the nuclei of the ordinarily stable elements. In order to penetrate the nuclei and to cause some kind of nuclear reaction, these projectiles must first penetrate the high potential barrier surrounding the bombarded nucleus. When Rutherford bombarded nitrogen nuclei by α particles and when Cockcroft and Walton bombarded lithium nuclei by artificially acclerated protons, the energy of the projectiles was always smaller than the heights of the potential barriers surrounding the nuclei in question. Thus the success of these experiments was entirely due to the quantum mechanical tunnel effect.

Questions

(28-1) **1.** What is the ratio of the number of neutrons to the number of protons in the nucleus of (a) Sc^{45}, (b) Cd^{112}, (c) Bi^{209}?

2. Calculate the neutron/proton ratio in the nuclei of (a) K^{39}, (b) Pd^{106}, (c) Xe^{131}.

(28-2) **3.** (a) What is the mass in gm of a nucleus of Ne^{10}? (b) What is its volume? (c) What is its radius?

4. Calculate, for a Ca^{40} nucleus: (a) its mass in gm, (b) its volume, (c) its radius.

5. Consider a mountain 1000 m high (average), about 5 km long, and 2 km wide; made of rock which has a density of 3 gm/cm^3. If all the atomic nuclei in this mountain were gathered in a single sphere, what would be its radius? (For this approximation, completely neglect the mass of the atomic electrons.)

6. The mass of the earth is 6×10^{27} gm. If all the nuclei of all the atoms in the earth were gathered into a single sphere, what would its radius be? (Neglect the mass of the atomic electrons.)

(28-3) **7.** (a) What is the binding energy of the alpha particle? (b) How many Mev of energy would be liberated if one were to make an alpha particle out of neutrons and protons? (Mass of $_2He^4 = 4.00260$ amu.)

8. The mass (atomic weight) of Li^7 is 7.01601 amu. (a) What is the binding energy of Li^7? (b) How much energy in Mev would be liberated if one could make a Li^7 nucleus out of neutrons and protons?

9. What is the binding energy per nucleon of He^4? (See Question 7.)

10. What is the binding energy per nucleon of Li^7? (See Question 8.)

11. What is the binding energy per nucleon of Fe^{56}? Give answer in amu and Mev.

12. What is the binding energy per nucleon of Zn^{64}? (mass = 63.9291.) Give answer in amu and Mev.

13. What is the binding energy per nucleon of U^{238}? (mass = 238.0508.)

14. What is the binding energy per nucleon of Pb^{208}? (mass = 207.9766.)

(28-4) 15. $Be^9 + He^4 \longrightarrow C^{12} + n$. (For Questions 15 through 20: (a) What is the mass (or energy) difference between the two sides of the reaction equation, in amu and in Mev? (b) Does this difference represent an excess of energy which the bombarding particle must have in order to conserve energy, or an excess of energy acquired by the product particles? Following are some atomic masses not given in the text: $Be^9 = 9.01219$, $B^{10} = 10.01294$, $B^{12} = 12.01437$, $Ne^{20} = 19.99244$, $Na^{23} = 22.98977$, $S^{32} = 31.97207$, $Cl^{35} = 34.96885$.)

16. $N^{14} + He^4 \longrightarrow H^1 + O^{17}$. (See Question 15.)

17. $S^{32} + He^4 \longrightarrow H^1 + Cl^{35}$. (See Question 15.)

18. $B^{10} + n \longrightarrow Li^7 + He^4$. (See Question 15.)

19. $H + Na^{23} \longrightarrow Ne^{20} + He^4$. (See Question 15.)

20. $Be^9 + He^4 \longrightarrow H^1 + B^{12}$. (See Question 15.)

(28-5) 21. (a) Consider the hypothetical fission of a nucleus of mass number 140 into two fragments of 60 and 80, respectively. From Fig. 28-2, estimate the total binding energy of the original particle, and of its two fragments. Does this fission release energy, or must energy be supplied? In either case, how much energy? (b) Repeat your calculations for the fission of a nucleus of mass number 80 into fragments of 40 and 40.

22. Follow the procedure of Question 21 for the fission of a nucleus of mass number (a) $160 \longrightarrow 100 + 60$. (b) $60 \longrightarrow 25 + 35$.

23. Follow the procedure of Question 21 for the fusion of (a) two nuclei each of mass number 10 into a single nucleus of mass number 20. (b) $40 + 60 \longrightarrow 100$.

24. Follow the procedure of Question 21 for the fusions: (a) $30 + 40 \longrightarrow 70$, (b) $50 + 60 \longrightarrow 110$.

chapter / twenty-nine

Large-scale Nuclear Reactions

29-1 Discovery of Fission

Neutrons are the ideal projectiles for nuclear bombardment because they have no electrical charge and thus suffer no repulsion in their approach to atomic nuclei. Following the discovery of neutrons, many new types of artificial nuclear transformations have been investigated. The impact of a neutron may result in the ejection of a proton or of an α particle, as in the following reactions:

$$_7N^{14} + {_0n^1} \longrightarrow {_6C^{14}} + {_1H^1}$$

$$_7N^{14} + {_0n^1} \longrightarrow {_5B^{11}} + {_2He^4}.$$

In some reactions, the incident neutron can eject another neutron without being captured itself:

$$_6C^{12} + {_0n^1} \longrightarrow {_6C^{11}} + 2{_0n^1}$$

whereas, in other cases, the incident neutron can be captured by the nucleus with the release of excess energy in the form of a γ photon. The latter process, known as the ***radiative capture*** of neutrons, is of particular

FIG. 29–1 Dr. Otto Hahn, who discovered that a uranium nucleus breaks up into two fragments when subjected to bombardment by neutrons.

importance for heavy nuclear targets, since in this case the ejection of protons and α particles is strongly hindered by the "outgoing" potential barrier surrounding the nucleus. The radiative capture of the neutron leads to the formation of a heavier isotope of the bombarded element. Sometimes these isotopes are stable, so that no further nuclear transformation takes place:

$$ {}_8O^{16} + {}_0n^1 \longrightarrow {}_8O^{17} + \gamma $$

whereas, in some other reactions, the radiative capture of a neutron leads to a β emission:

$$ {}_{47}Ag^{109} + {}_0n^1 \longrightarrow {}_{47}Ag^{110} + \gamma $$
$$ {}_{47}Ag^{110} \longrightarrow {}_{48}Cd^{110} + e^- $$

which is necessary to reestablish a stable neutron-to-proton ratio.

In the year 1939, a German radio chemist, Otto Hahn, with his co-worker Fritz Strassman studied the effect of neutron bombardment of uranium atoms, expecting to observe the formation of uranium isotopes with atomic weights higher than that of ordinary uranium. To his great surprise (Fig. 29-1), Hahn found that the sample of uranium

bombarded by neutrons contained radioactive atoms of a much lighter element, barium. The mystery of this discovery was soon cleared up by two German physicists, Lise Meitner and Otto Frisch, who suggested that in Hahn's and Strassman's experiments the nuclei of U^{238} were split by incident neutrons into two nearly equal parts:

$$ {}_{92}U^{238} + {}_{0}n^{1} \longrightarrow {}_{56}Ba^{145} + {}_{36}Kr^{94}. $$

Since the barium and krypton atoms produced in this process possessed excess neutrons, as compared with ordinary stable atoms of the same atomic weight (${}_{60}Nd^{145}$ and ${}_{40}Zr^{94}$), these so-called *fission products* emitted negative electrons, making them strongly radioactive. Frisch and Meitner's interpretation of Hahn and Strassman's experimental finding as the splitting of the uranium nucleus into two nearly equal parts opened new vistas in the field of nuclear physics. Instead of just "chipping off" small pieces of the bombarded nucleus, as was the case in all previous experiments, here was a real breakup of the central body of the atom, the *fission* of a large droplet of the nuclear fluid into two half-size droplets. Instead of just the few million electron volts of energy observed in previous experiments on artificial nuclear transformations, uranium fission liberates 200 Mev per atom!

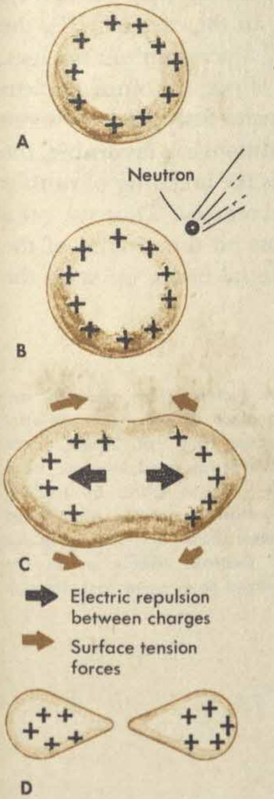


FIG. 29–2 What happens when a massive nucleus breaks up in the fission process.

Detailed theoretical studies of the process of nuclear fission were carried out by Niels Bohr and John Wheeler (1911–) and published in the September 1939 issue of the *Physical Review*. This was the first and last comprehensive article on the theory of nuclear fission that appeared as open literature before the security curtain was drawn tight on that subject. According to Bohr and Wheeler, the fission of heavy nuclei resulting from the impact of a neutron is a resolution of a conflict between the opposing tendencies of nuclear (attractive) and Coulomb (repulsive) forces acting in the atomic nucleus. If we imagine an atomic nucleus to be an electrically charged droplet of nuclear fluid (Fig. 29-2A), we find that any excitation (Fig. 29-2B) transforming its initial spherical shape into a more or less elongated ellipsoid (Fig. 29-2C) causes two kinds of forces to act on it:

1. The forces of nuclear surface tension attempting to restore the nucleus to its original spherical shape (the same as in the case of a deformed spherical droplet of water or mercury).

2. Coulomb repulsive forces between the electric charges on the opposite ends of the ellipsoid attempting to break up the nucleus into two halves.

In the previous chapter, we saw that the nuclear fluid model leads us to the conclusion that for the lighter nuclei the surface-tension forces have the upper hand, but that for the heavier nuclei the electric forces become more and more important. Thus we would expect that in the

case of very heavy nuclei the comparatively small deformation caused by a neutron impact might result in a breakup (fission) of the original nuclear droplet into two halves (Fig. 29-2D).

29-2 Fission Neutrons

In spite of the fact that each of the two fragments produced in the fission of a uranium nucleus carries about 100 Mev of energy, these fragments are quite ineffective in producing further fission processes; this is due to the fact that the fission fragments carry a very high electric charge and are consequently strongly repelled by the other uranium nuclei with which they may collide. Thus the discovery of uranium fission would not contribute anything to the problem of the large-scale liberation of nuclear energy if it were not for a secondary process that was found to accompany nuclear fission. It was discovered that in addition to the two large fragments of the original nucleus, there are always several extra neutrons emitted in the breakup. In the case of U^{235}, the average number of "fission neutrons" formed is 2.5 per uranium nucleus. These fission neutrons formed in the breakup of one uranium nucleus may collide with the surrounding uranium nuclei and produce more fission and still more fission neutrons. If the conditions are favorable, the breeding of fission neutrons goes *crescendo* as does the breeding of rabbits on a rabbit farm or of fruit flies in a genetics laboratory. Thus we get a *branching chain reaction* (Fig. 29-3), and in almost no time many of the nuclei of uranium in a given piece of this material break up with the liberation of a tremendous amount of energy.

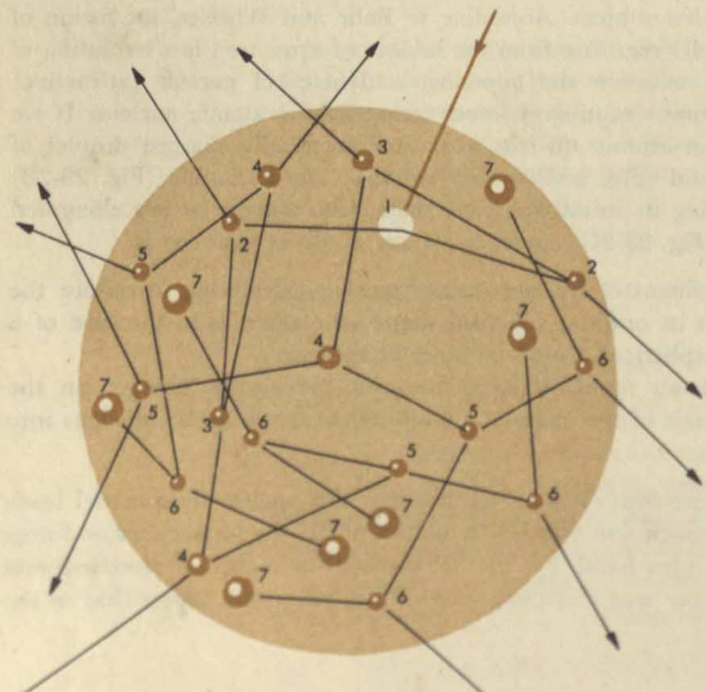


FIG. 29-3 A nuclear chain reaction developing in a piece of fissionable material with a "branching ratio" of 2—i.e., each neutron-induced fission produces two more neutrons. The reaction starts at 1 by a single neutron from the outside. After seven successive generations in this particular piece, seven neutrons remain inside the piece, and eleven have been lost through the surface.

29-3 Fissionable Uranium-235

Natural uranium represents a mixture of two isotopes, U^{238} and U^{235}, that are present in the relative amounts of 99.3 and 0.7 percent respectively. The study of these two isotopes under the influence of neutron bombardment had shown that the rarer isotope U^{235} is much more fissionable than the more abundant U^{238}. Indeed, whereas U^{238} nuclei do not break up unless the bombarding neutron has energy above 1 2 Mev, U^{235} nuclei can be broken up by neutrons moving with much smaller velocities, and, in fact, the probability of breaking up increases with decreasing velocity of incident neutrons.

In the range between the high energies needed to fission U^{238} and the very low, or "thermal," energies favorable for fissioning U^{235}, neutrons are absorbed by U^{238} without causing the latter to fission. Thus the strong dilution of the active U^{235} isotope by the inactive U^{238} makes natural uranium as useless for carrying out nuclear chain reactions as wet logs are for building a campfire. Indeed, most of the fission neutrons ejected in the breakup of U^{235} nuclei in natural uranium will be captured by the much more abundant U^{238} nuclei and thus will be taken out of the game.

Accordingly, in the early stages of nuclear energy development (the "Manhattan Project"), much effort was spent on the separation of the active U^{235} from the inactive U^{238}. Since the isotopes of a given element possess identical chemical properties, ordinary chemical separation methods could not be used in this case. The problem was finally solved by the development of the "diffusion-separation" method, which was based on the fact that the lighter atoms of U^{235} (and their various chemical compounds) diffuse faster through tiny openings in a porous membrane than do the heavier U^{238} atoms, and large amounts of fissionable uranium-235 were obtained in this way.

29-4 The Fermi Pile and Plutonium

A good boy scout is supposed to be able to build a campfire even if the wood is soaking wet. This role of a good boy scout in the nuclear energy project was played by the Italian-American physicist Enrico Fermi (Fig. 29-4), who actually made the wet uranium logs burn. He was able to do so by utilizing the fact mentioned above, that the effectiveness of fission neutrons in producing the fission of U^{235} nuclei increases when they are slowed down. If such slowing down of fission neutrons could be achieved, the presence of inactive U^{238} would not make much difference, because very-low-energy, slow neutrons are not absorbed by U^{238} to any appreciable extent. To slow down the fast fission neutrons it was necessary to use a *moderator*—i.e., some material from whose atoms the fast neutrons could bounce harmlessly and lose their energy. From considerations of conservation of momentum and energy, it can be shown that when a particle collides with another particle much less massive

FIG. 29–4. Enrico Fermi looking at the plaque on the wall of the University of Chicago Stadium (Stagg Field) which reads: "On December 2, 1942, man achieved here the first self-sustaining chain reaction and thereby initiated the controlled release of nuclear energy."

than itself, it is slowed down very little and loses scarcely any energy (recall the lack of effect that electrons have in deflecting α particles). At the other extreme, if a particle collides with another particle much more massive than itself, it bounces back with a speed and energy that are little changed. To be most effective, then, in helping the neutron lose energy by collision, the moderator atoms should be light atoms, comparable in size to the neutron, and should not absorb neutrons. It was decided to surround the pieces of uranium by carbon for a moderator, in the form of very pure graphite.

Of course, from the purely mechanical point of view, ordinary hydrogen (or some hydrogen-rich material, such as plain water) would be the best moderator. Its nuclear protons, having almost the same mass as the neutrons, would absorb the maximum possible energy in each collision and would be most effective in slowing the neutrons. Unfortunately, however, some of the collisions would result in the absorption of the neutron, and its loss from the fission chain:

$$_0n^1 + {}_1H^1 \longrightarrow {}_1H^2.$$

This heavy isotope of hydrogen, $_1H^2$, is such a common material in nuclear research that it has even been given its own name—*deuterium*—and its own chemical symbol D. Its nucleus, a proton and a neutron, is called a *deuteron*. About 1 H atom in 7000 is actually a D atom, so deuterium is relatively plentiful. "Heavy water," D_2O, is a fine neutron moderator, and can be separated from H_2O by a long and tedious series of repeated electrolysis or distillation; because of the smaller mass and greater mobility of its molecules, H_2O distills and electrolyzes a little more readily than D_2O.

Scientists in the nuclear project in the United States, however, decided to use the more immediately available element carbon as the moderator. A large "pile" of graphite bricks with small pieces of natural uranium included in the structure (Fig. 29-5 and 29-6) was constructed in great secrecy under the grandstand of the University of Chicago's Stagg Field, and on December 2, 1942, Professor Compton phoned to his colleague, Professor Conant of Harvard, the guarded message: "The Italian navigator has landed. The natives are friendly." This was quite correctly interpreted to mean: "Fermi's pile works successfully. The first successful nuclear chain reaction has been achieved."

In the pile, the fission chain reaction could be maintained in natural uranium, but the natural uranium was so highly diluted by carbon that high efficiency in energy production could not be achieved. Owing to the presence of inactive U^{238}, the chain reaction in the pile could not

FIG. 29–5 The only photograph (taken in November, 1942) made during the construction of the first nuclear reactor. Here the nineteenth layer (almost covered up) is formed by graphite blocks studded with spheres of uranium metal and uranium oxide. The even-numbered layers consisted only of graphite blocks without uranium. The pile became critical on December 2, 1942, after 57 layers had been laid down,

FIG. 29–6 A drawing of the first nuclear reactor, constructed in a squash court under the west stands of Chicago University's Stagg Field in 1942.

Courtesy Argonne National Laboratory.

Courtesy U.S. Army Engineers.

possibly develop into an efficient explosion, nor could it be very useful as a power source. So what good was the pile, except for demonstrating the purely scientific principle of the possibility of a self-maintaining nuclear reaction? Of course, the demonstration of a purely scientific principle is always of very great importance, but the pile was built at great expense in the midst of a perilous war when all expenditures were supposed to be judged on the basis of their military usefulness.

Fermi's reactor stood this test. Although the energy released in the fission of U^{235} nuclei could not be utilized and was literally sent down the drain by means of the water-cooling system, a new fissionable element was produced inside the pile during its operation. The neutrons that were not used in the maintenance of the chain reaction in U^{235} nuclei were captured by U^{238} nuclei, producing a heavier isotope:

$$_{92}U^{238} + {_0n^1} \longrightarrow {_{92}U^{239}} + \gamma.$$

Having an excess of neutrons, the nuclei of $_{92}U^{239}$ underwent two successive β transformations, giving rise to elements with atomic numbers 93 and 94. These two elements, which do not exist in nature but had been produced artificially by human genius, were given the names *neptunium* and *plutonium*. The reactions following the neutron capture by U^{238} can be written as follows:

$$_{92}U^{239} \longrightarrow {_{93}Np^{239}} + e^-$$

$$_{93}Np^{239} \longrightarrow {_{94}Pu^{239}} + e^-.$$

Being chemically different from uranium, the plutonium produced can be separated and purified with much less effort than it takes to separate the light uranium isotope from the heavy one, and this element turned out to be even more fissionable than U^{235}. In fact, whereas U^{235} gives rise to 2.5 fission neutrons, the corresponding figure for Pu^{239} is 2.7 fission neutrons.

29-5 Critical Size

When a single fission process occurs inside a given piece of pure U^{235} or Pu^{239}, several fission neutrons are ejected from the point where the nuclear breakup takes place. The average distance a fission neutron must travel through the material before it is slowed down to the point where it can effectively cause a fission reaction is about 10 cm, so that if the size of the sample in question is less than that, most of the fission neutrons will escape through the surface and fly away before they have a chance to cause another fission and produce more neutrons. Thus no progressive chain reaction can develop if the piece of fissionable material is too small. Going to larger and larger sizes, more and more fission neutrons produced in the interior have a chance to produce another fission by colliding with a nucleus before they escape through the sur-

face; and for very large pieces, only a small fraction of the neutrons produced in them has a chance to reach the surface before fissioning one of the nuclei. The size of the sample of a given fissionable material for which the percentage of neutrons giving rise to subsequent fission processes is high enough to secure a progressive chain reaction is known as the *critical size* for that particular material. Since the number of neutrons per fission is larger in the case of plutonium than in the case of uranium-235, the critical size for plutonium is smaller than that for uranium-235, because the former can afford larger losses of neutrons through its surface.

29-6 Nuclear Reactors

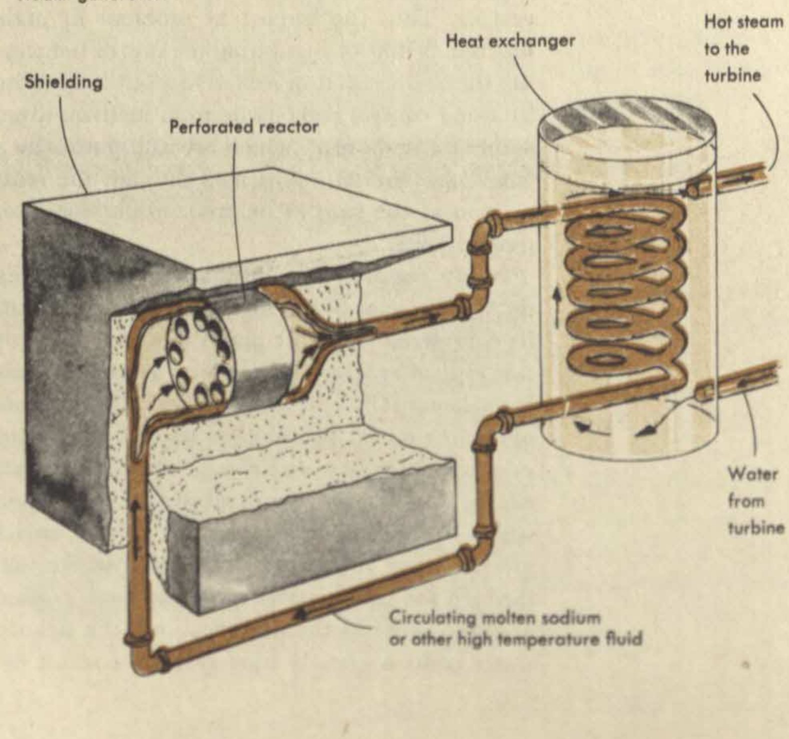
As we have just seen, a piece of fissionable material smaller than the critical size is unable to carry on a nuclear chain reaction. If the size is *exactly critical*, the number of neutrons produced in each generation is the same as that produced in the previous one, resulting in steady nuclear-energy liberation. The original Fermi pile and its later modifications maintained nuclear reactions at the critical level. It must be mentioned in this connection that the conditions of "criticality" are extremely unstable: a small deviation in one direction will result in the rapid extinction of fission neutrons and the cutoff of the nuclear chain reaction, whereas a deviation in the other direction will lead to a rapid multiplication of the fission neutrons and the melting of part of the reactor. Thus the important problem in maintaining a steady chain reaction is that of regulating the rate of neutron production and of keeping the chain reaction from dying out or running away. This is achieved by using control rods made from neutron-absorbing materials (such as cadmium or boron) which are automatically pushed in or pulled out from narrow channels drilled through the reacting fissionable material as soon as the rate of neutron production exceeds or drops below the desired level.

Many reactors are unsuitable for purposes of nuclear-power production because of the high dilution of uranium by carbon; they should be considered rather as plants in which plutonium is produced. Because this type of reactor can produce more fissionable material (Pu^{239}) than it consumes (U^{235}), it is sometimes called a *breeder reactor*. For the purpose of nuclear power production, we can use controlled nuclear chain reactions in relatively pure fissionable materials, such as U^{235} or Pu^{239}, which can be run at high temperatures. In the "swimming pool" reactor, in which cylindrical containers filled with enriched fissionable material are placed at the bottom of a large water tank, the water circulating through the tank carries away the heat produced in the fission process and also protects the observer from the deadly nuclear radiation. The water emits a ghostly blue glow as a result of what is called *Cherenkoff*

radiation, after the Russian physicist who first analyzed the phenomenon in 1934. (For this and related work, Cherenkoff received the 1958 Nobel Prize in Physics.) Cherenkoff radiation is produced by two separate effects that we have previously become acquainted with: the Compton effect and shock waves. We have spoken of the Compton effect before in terms of the scattering of an X-ray photon and its consequent loss of energy because of collision with an electron. In the reactor tank, high-energy γ photons interacting with electrons send the electrons flying off at speeds greater than the speed of light *in the water*. The result is similar to the bow wave of a ship traveling through water faster than the surface ripples can spread out: a shock wave (a bow wave for the ship) is formed. (See Fig. 9-2.) Analysis beyond the scope of this book shows that this electromagnetic shock wave will give rise to the blue Cherenkoff radiation.

Many reactors are now being built all over the world for the purpose of generating electrical power. Unfortunately, in all of these the released nuclear energy can be used only as a source of heat. The generators are driven by the same steam turbines to be found in a coal-burning installation (Fig. 29-7). So nuclear generating plants must operate at the same

FIG. 29-7 Schematic diagram of a closed-cycle reactor, in which a sealed-in high-temperature fluid carries heat from the reactor to the steam generator.



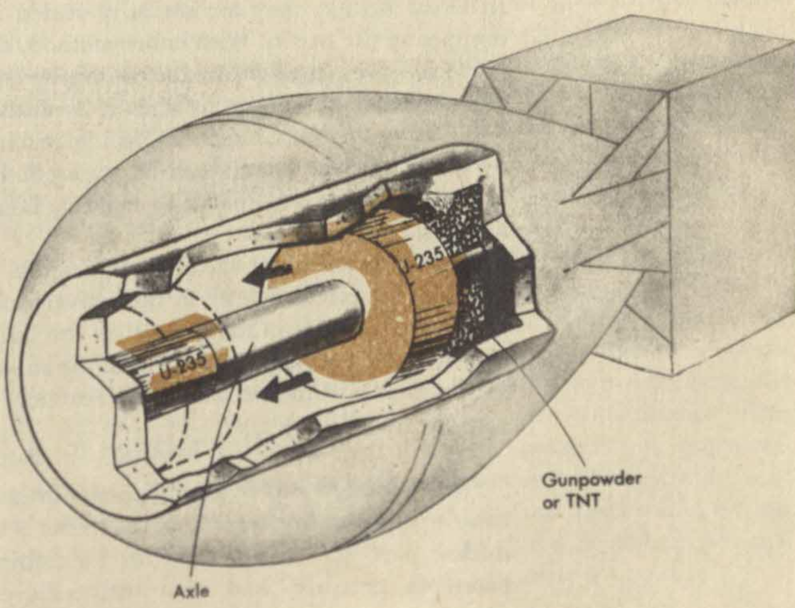
temperatures as conventional plants, and are limited by the same low thermal efficiencies.

Much research is being carried on in an attempt to convert nuclear energy directly into electrical energy without using the inefficient heat engine as an intermediate step. A commercially practical system has not yet been developed.

29-7 Fission Bombs

If a piece of nearly pure fissionable material exceeds the critical mass, the number of fission neutrons and the rate of energy production will increase rapidly with time and the process will acquire an explosive nature. The principle of the fission bomb, or the "atomic bomb" as it is commonly called, consists of building up (assembling) a mass of fissionable material of more than the critical size in such a short time period that the nuclear-energy liberation that starts at the beginning of the assembly period does not develop to any important degree before the assembly job is finished. This can be accomplished in a simple way by inserting one subcritical piece of fissionable material into another subcritical piece as indicated in Fig. 29-8. In order to perform the assembly

FIG. 29-8 The principle of the "gun-type" atomic bomb. A conventional explosive drives a U-235 or plutonium cylinder along the axle until it surrounds another piece of fissionable material forming the far end of the axle, thereby producing a single fissionable mass of more than critical size.



process fast enough, we must shoot the inserted piece at high speed with a conventional chemical explosive, which earned for this assembly method the name of "gun gadget." There are also other more ingenious methods of bringing a given amount of fissionable material to supercritical size.

The energy liberation in the explosion of nuclear bombs is measured, according to established convention, in units known as "kilotons" and "megatons," which refer to the weight of TNT (ordinary high explosive) that liberates the same amount of energy. One kiloton, i.e., the energy liberated in the explosion of 1000 tons of TNT, equals about 5×10^{19} ergs, or 5×10^{12} joules.

29-8 Fusion Reactors

Fission reactors are economical and efficient producers of energy, but they do have certain defects. There have already been a few occasions on which the automatic controls have failed to function properly, so that the reactor has produced heat and energy faster than it could be used. Although there is no danger of a bomblike explosion, these accidents have resulted in contaminating the reactor plant and the surrounding countryside with radioactive material.

Even if no such mishap occurs (and the new designs are much safer than the old), all fission reactors present the serious problem of disposing of their highly radioactive fission products. In the United States these dangerous materials are concentrated, sealed in stainless steel drums, and encased in concrete in the dry caves of our desert regions. In Great Britain they are similarly sealed in steel and concrete and dropped in the ocean. With either method, the hope is that many half-lives will have passed before the containers corrode, so that the material will no longer be dangerous when it is released.

But as we saw in Chapter 28, fission reactors are not the only means by which nuclear energy can be released. Fusion reactions, in which two small nuclei are joined to make a larger one, also release large amounts of energy.

Nuclear fusion is not one of man's original inventions; it has been the principal source of energy in the universe for many billions of years. As we shall discuss in more detail in the last chapter, nearly all of the enormous outpouring of energy from the sun and other stars comes from the simplest possible nuclear fusion reaction:

$$4\,{}_1H^1 \longrightarrow {}_2He^4 + 2e^+ + 26 \text{ Mev}.$$

Simple though this reaction appears to be, it is not one that we can cause to occur easily here on earth. In the center of the sun the reacting nuclear particles are held together for millions of years, if need be, at enormous pressure, and at a temperature of about 20,000,000°K.

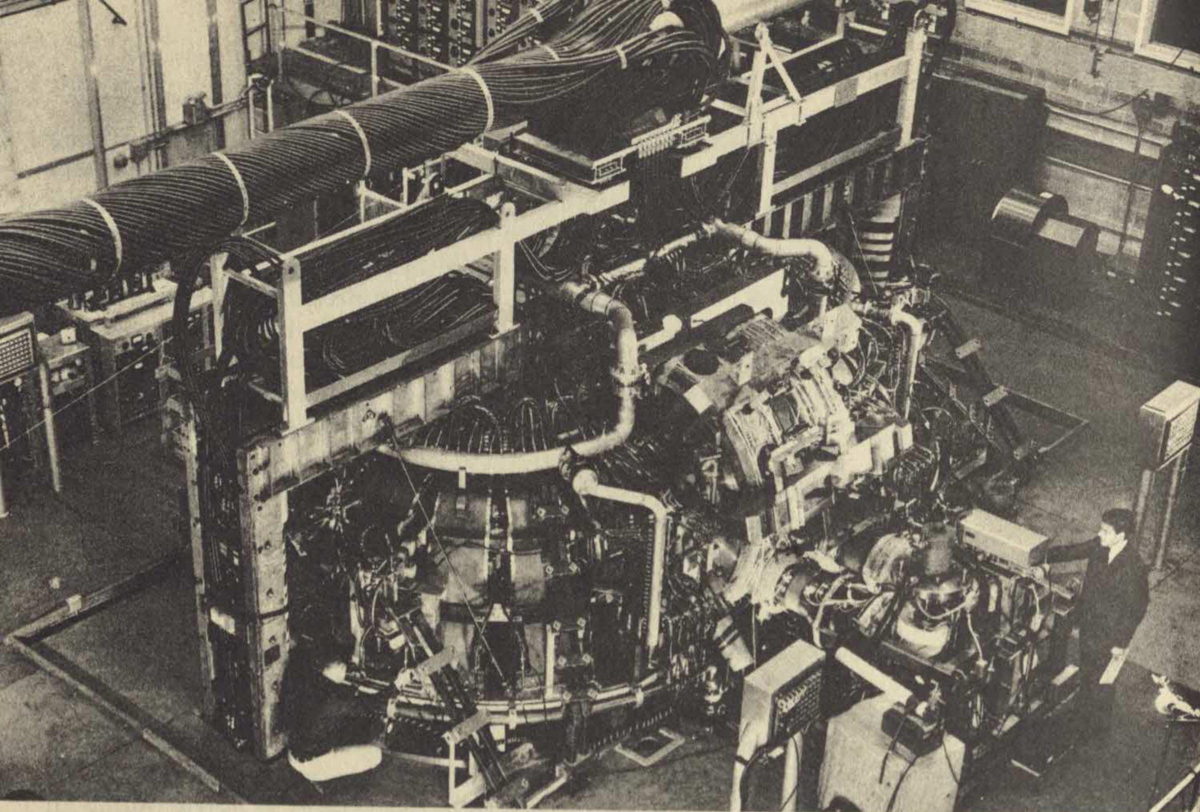
Such high temperature is necessary, because the colliding nuclei must have enough energy to overcome their electrostatic repulsion. One place where high pressure and multimillion-degree temperatures are available is in the immediate vicinity of an exploding nuclear fission bomb. For this reason the so-called "hydrogen bomb" (actually, it is rumored, now made largely of lithium hydride, LiH) is extremely simple, in principle. Essentially, all that needs to be done is to surround a fission bomb with a layer of material of low atomic number. When the fission bomb explodes, it provides an environment in which the light nuclei are forced to come together and release still more energy of nuclear fusion.

Such a procedure, however, is far from satisfactory for operating a power plant. Intensive work is being carried out by all countries which have the necessary financial and scientific resources, in an effort to find some practical way of maintaining a steady, controllable flow of energy from nuclear fusion reactions.

In fusion experiments in which beams of particles from accelerators are directed against light-element targets, the yield of the desired reactions is very low. The charged particles, such as protons or α particles, rapidly lose their energy in the ionization of the target material, and only a very few of them (about 0.01 percent) have a chance to collide with another nucleus before having spent their energy in merely tearing off atomic electrons. Such experiments are extremely valuable for the study of nuclear properties, but are worthless for the purpose of large-scale nuclear energy liberation.

In man's search for workable fusion reactions (often called *thermonuclear reactions*), the fuel used is generally composed of the easier-reacting heavy isotopes of hydrogen—deuterium, which is plentiful in the natural world, and tritium, which is so rare it must itself be produced by nuclear reaction processes. It is relatively easy to ionize these hydrogen isotopes and to use something like an electron gun to accelerate the resulting bare nuclei to energies corresponding to many millions, or even billions, of degrees.

The problem is how to confine this *plasma*—this gas of ionized particles—until enough collisions take place to produce a profitable amount of nuclear fusion energy. A material container is out of the question, because the plasma would at once be cooled off to the temperature of the confining walls. Since charged particles are subjected to strong forces when they move rapidly through a magnetic field, the most common experimental approach has been to confine the particles by ingeniously designed magnetic fields. The problem is difficult, though, and has not yet been solved. Figure 29-9 shows one heretofore unsuccessful attempt. However, most scientists agree that if something is possible in principle, human ingenuity will eventually make it possible in practice.



Courtesy Dr. Lyman Spitzer, Jr.

FIG. 29-9 The Model C "Stellarator" at Princeton University.

When a successful fusion reactor is finally designed, it will have many advantages over the fission reactor. No radioactive fission fragments will be produced to present the hazards of contamination. The disposition of radioactive wastes will no longer be a problem. The ores of uranium and thorium (which can also be used to produce a fissionable isotope in a breeder reactor) are relatively scarce, and must be prospected for, mined, and refined at great expense. Deuterium, however, can easily be extracted from sea water, of which we have an enormous supply.

The total volume of the oceans is 330 million cubic miles, and each cubic mile of water contains about 500 million metric tons of hydrogen. This gives us a total of $3.3 \times 10^8 \times 5 \times 10^8 = 1.6 \times 10^{17}$ tons of hydrogen. Approximately 0.02 percent of this is deuterium, so our supply of this fusion fuel is $1.6 \times 10^{17} \times 2 \times 10^{-4} = 3.2 \times 10^{13}$ metric tons. It is known that the fusion of 1 gm of deuterium will produce 7.5×10^{10} joules of energy, so our total available supply from the oceans amounts to $7.5 \times 10^{10} \times 3.2 \times 10^{13} \times 10^6 = 2.4 \times 10^{30}$ joules.

If this figure does not seem impressively large, consider that the world's total use of energy for industry, transportation, heating, etc., amounts

to about 2×10^{18} joules per year at present. Thus the deuterium in the ocean could supply energy for all our needs, at the present rate, for at least 10^{12} years. This is more than a hundred times longer than the sun will continue to shine at a rate that will permit life to continue on the earth.

Questions

(29-1)

1. Is the neutron/proton ratio increased or decreased when a radioactive nucleus emits (a) an electron? (b) a positron?

2. What happens to the neutron/proton ratio when a nucleus (a) emits an alpha particle? (b) emits a photon?

3. Explain why we should expect fission fragments to have a high n/p ratio, in comparison with stable nuclei of the same mass number.

4. Will radioactive fission fragments be more likely to decay by emission of electrons, or of α particles?

5. Assume a U^{235} nucleus fissions into two fragments of mass numbers 100 and 135. Refer to Fig. 28-2, and calculate the approximate total energy release.

6. A nucleus of Pa^{239} fissions into fragments of mass numbers 80 and 159. From the data of Fig. 28-2, calculate the approximate total energy release.

(29-3)

7. In the diffusion separation of U^{235} and U^{238}, which depends on the relative speeds of the molecules involved, the gaseous compound UF_6 (uranium hexafluoride) was used. At any given temperature, (a) which molecules have the higher speed, $U^{238}F_6$ or $U^{235}F_6$? (b) What is the ratio of their speeds?

8. To a lesser extent devices similar to mass spectrographs were also used for uranium separation, depending on the e/m ratios of singly charged ions of the two principal uranium isotopes. (a) Which has the larger e/m; U^{235+} or U^{238+}? (b) What is the ratio of their e/m's?

(29-4)

9. Consider a direct, head-on, perfectly elastic collision between a moving particle of mass m and a stationary particle of mass $10\ m$. (a) What fraction of its kinetic energy would m lose in this collision? How many such collisions would be needed to reduce the KE of m to less than 5 percent of its initial value?

10. Same as Question 9, except that the target particle has a mass of only $2\ m$.

(29-7)

11. The earliest nuclear fission bombs were rated at 20 kilotons. (a) How many ergs of nuclear energy were released? (b) How many Mev? (c) How many U^{235} or Pu^{239} atoms must have fissioned? (d) How many grams of uranium are actually fissioned in a 20-kiloton bomb?

12. How long would it take a small 1 megawatt generating plant to produce as much energy as a single 50-kiloton bomb?

(29-8)

13. Particles at room temperature (about 300°K) have thermal energies of approximately 0.025 ev. What is the thermal energy of atomic and nuclear particles at the center of the sun?

14. In the last chapters we have become quite nonchalant about particles with energies of 1 Mev. About what theoretical temperature would be required to give particles this much thermal energy? (See Question 13.)

15. About how often would one need to explode a 1-megaton bomb in order to produce energy at the rate the world's population is at present using it?

chapter / thirty

Mystery Particles

30-1 The Positron: Antielectron

In the early days of nuclear physics, the materials considered necessary to make the universe comprised a very short list: electrons, protons, and photons. It was still a relatively simple world when the neutron was added to this list in 1932. But three years earlier, in 1929, the purely theoretical picture had been complicated by a British physicist, P. A. M. Dirac (Fig. 30-1), who was busy trying to reconcile the basic principles of the quantum theory with those of Einstein's theory of relativity. On the basis of very abstract theoretical considerations, Dirac came to the conclusion that in addition to the "ordinary" electrons which revolve around atomic nuclei or fly through vacuum tubes, there must also exist an incalculable multitude of "extraordinary" electrons distributed uniformly throughout what one usually calls empty space. Although, according to Dirac's views, each unit volume of vacuum is packed to capacity with these extraordinary electrons, their presence escapes any possible experimental detection. The "ordinary" electrons studied by physicists and utilized by radio engineers are those few excess

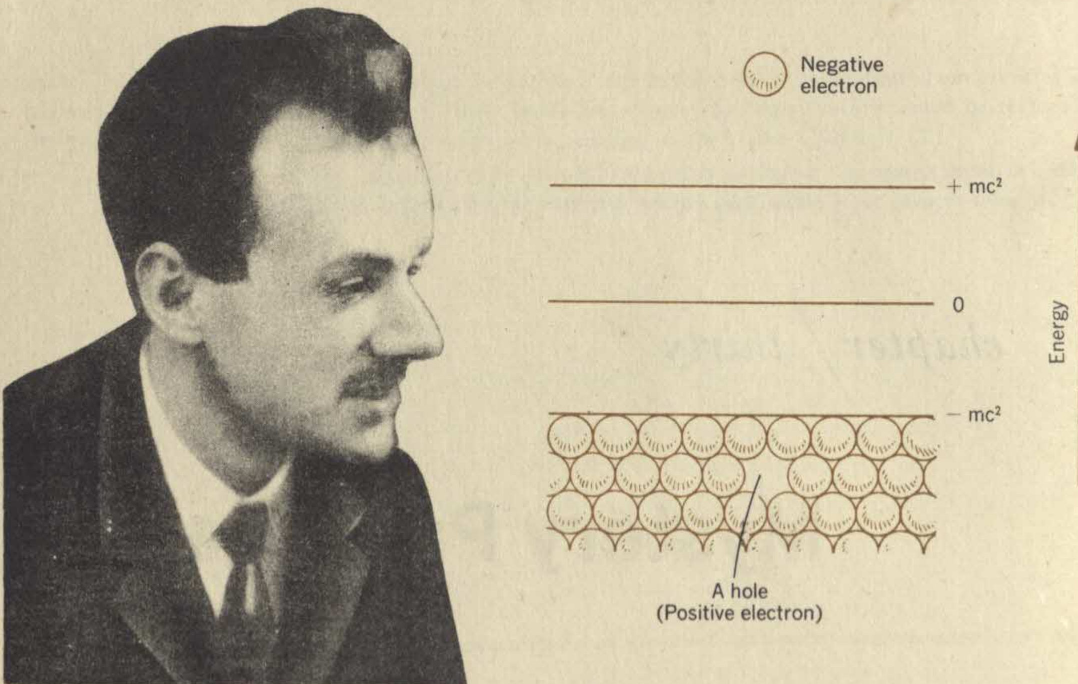


FIG. 30–1 Dr. P. A. M. Dirac, who conceived the idea that "empty" space is actually tightly packed with electrons of negative mass that are inaccessible to any physical observation. We can observe an electron only when it is raised into the region of positive energy (above, on the right). The removal of an electron from the continuous distribution leaves a "hole" (below, on the right), which represents a positive electron.

particles that cause an overflow of Dirac's ocean of "extraordinary" particles (Fig. 30-1, right), and they thus can be observed individually. If there is no such overflow, nothing can be observed, and we call the space empty.

In addition to having the property of not being observable by any physical means, these extraordinary electrons possess, according to Dirac, a *negative mass*, and of course a negative $E = mc^2$ energy, and (if they are moving) a negative kinetic energy.

Because of their uniform distribution, the extraordinary electrons forming Dirac's ocean are invisible to observation, but what happens if one of these particles is absent, leaving in its place an empty hole? (See Fig. 30-1.) This hole in the uniform distribution of particles of negative charge and mass represents the *lack of a negative charge*, which is equivalent to the *presence of a positive charge*; and the *lack of a negative mass*, which is the same as having our ordinary *positive mass*. Thus the electrical instruments used in our physical laboratories would register

this hole as a positively charged particle with the same numerical value of charge as an ordinary electron, but with the opposite sign, and the same ordinary mass as an electron. This is similar to the notion of holes in the uniform distribution of electrons in semiconductors, which led to a successful explanation of their properties. But, whereas in that case the notion of a hole can be readily visualized on the basis of an ordinary picture of the electric nature of matter, Dirac's holes belong to a much more abstract physical picture. His ocean surrounding us on all sides and extending into infinity in all directions remains unobservable to us. In a sense, Dirac's theory brings us back to the old-fashioned idea of the "all-penetrating ether," but in an entirely new fashion.

During the first few years after its publication, Dirac's paper was subjected to a great deal of criticism. The criticism stopped abruptly in 1932, when an American physicist, Carl Anderson, confirmed by direct observation the existence of the new particles predicted by Dirac's theory. These particles, carrying a positive electron charge and a positive mass equal to that of the electron, are called *antielectrons*, *positive electrons*, or simply *positrons*.

30-2 Pair Production and Annihilation

From what has been said, we can conclude that in order to form a positron, we have to remove a negative electron from its place in Dirac's ocean. When this electron is removed from the uniform distribution of negative electric charge, it becomes observable as an ordinary negatively charged particle. Thus *the positive and negative electrons always must be formed in pairs*. We often call this process *the "creation" of an electron pair*, which is not quite correct because the pairs of electrons are not created from nothing but are formed at the expense of the energy spent in carrying out the process of their formation. According to Einstein's $E = mc^2$, the energy necessary to produce two electron masses is $E = 2 \times 9.11 \times 10^{-28} \times (3 \times 10^{10})^2 = 1.64 \times 10^{-6}$ erg, or 1.02 Mev (million electron-volts). Thus, if we irradiate matter with electromagnetic radiation whose photons have this much energy or more, we should be able to induce the formation of pairs of positive and negative electrons. The electron pairs discovered by Anderson were produced in atmospheric air, and also in metal plates placed in a detecting cloud chamber, by the high-energy γ radiation associated with the cosmic rays which fall on the earth from interstellar space (Fig. 30-2). After this discovery, physicists learned to produce electron pairs by irradiating different materials by high-energy γ rays.

The opposite of the "creation" of an electron pair is the "annihilation" of a positive electron in a collision with an ordinary negative electron. According to Dirac's picture, the annihilation process occurs when an ordinary negative electron, which moves "above the rim" of the com-

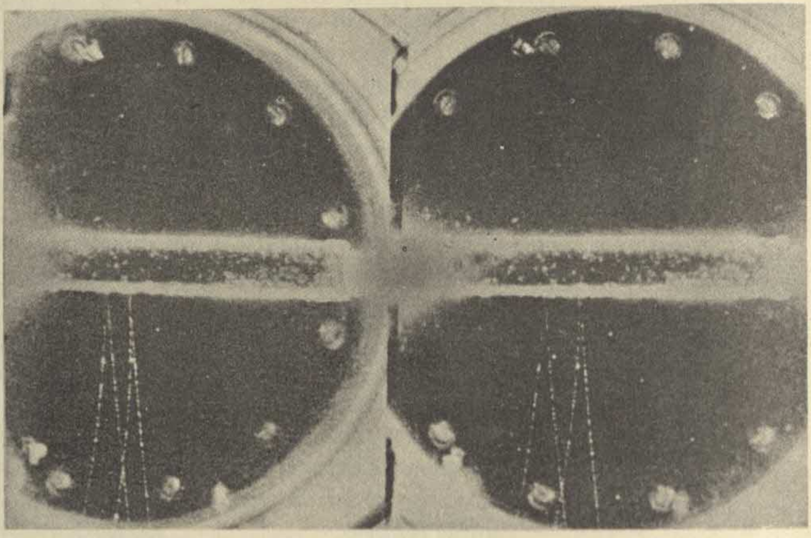


Photo by Dr. Carl Anderson, California Institute of Technology.

FIG 30–2 A cloud-chamber photograph of two electron pairs produced in a metal plate by a high-energy cosmic ray photon.

pletely filled Dirac's ocean, finds a "hole" in the distribution and falls into it. In this process the two individual particles disappear, giving rise to two photons of γ radiation with a total energy equivalent to the vanished mass, radiating from the place of encounter. We can readily see that *two* photons are necessary by considering conservation of momentum. Imagine the mutual annihilation of a slow-moving electron and a slow-moving positron, whose total momentum is almost zero. A single photon would have a large momentum in one direction; for the photon momentum to be zero, there must be two photons in opposite directions. Dirac's original theory of "holes" not only predicted the existence of positive electrons before their experimental discovery but also gave an excellent mathematical apparatus for calculating the probabilities of the formation of electron pairs under different circumstances, as well as the probability of their annihilation in casual encounter. All the predictions of this theory stand in perfect agreement with experimental evidence.

30-3 Antiprotons and Antineutrons

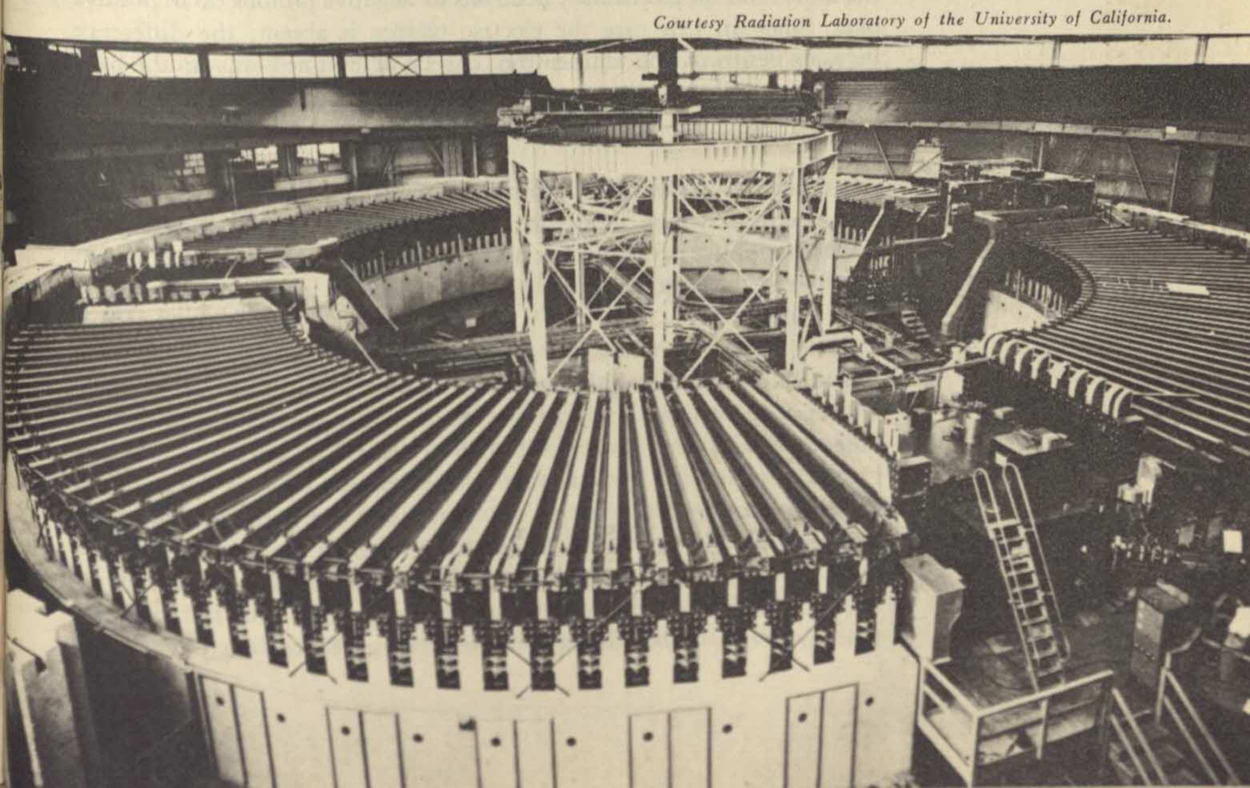
After the experimental confirmation of Dirac's theory of antielectrons, physicists were interested in finding the *antiprotons* that should be the particles of proton mass carrying a negative electric charge, i.e. *negative*

protons. Since a proton is 1840 times heavier than an electron, its formation would require a correspondingly higher input of energy. It was expected that a pair of negative and positive protons should be formed when matter is bombarded by atomic projectiles carrying not less than 4.4 Bev (billion electron-volts) of energy. With this task in mind, the Radiation Laboratory of the University of California at Berkeley and the Brookhaven National Laboratory on Long Island, New York, started construction of the gigantic particle accelerators—*Bevatron* on the West Coast (Fig. 30-3) and *Cosmatron* on the East Coast—that were supposed to speed up atomic projectiles to the energies necessary for the proton-pair production. The race was won by the West Coast physicists, who announced in October, 1955, that they had observed negative protons being ejected from targets bombarded by 6.2-Bev atomic projectiles.

The main difficulty in observing the negative protons formed in the bombarded target was that these protons were expected to be accompanied by tens of thousands of other particles also formed during the impact. Thus the negative protons had to be filtered out and separated from all the other accompanying particles. This was achieved by means of a complicated labyrinth formed by magnetic fields, narrow slits,

FIG. 30–3 The giant accelerator of the University of California, known as the Bevatron. This machine first enabled scientists to detect antiprotons and antineutrons.

Courtesy Radiation Laboratory of the University of California.



etc., through which only the particles possessing the expected properties of antiprotons could pass. When the swarm of particles coming from the target (located in the bombarding beam of the Bevatron) was passed through this labyrinth, the experimentalists were gratified to observe the expected particles coming out at a rate of about one every 6 minutes. As further tests showed, the particles were genuine negative protons formed in the bombarded target by the high-energy Bevatron beam. Their mass was found to have a value of 1840 electron masses, which is the mass of ordinary positive protons.

Just as the artificially produced positive electrons are annihilated in passing through ordinary matter containing a multitude of ordinary negative electrons, negative protons are expected to be annihilated by encountering positive protons in the atomic nuclei with which they collide. Since the energy involved in the process of proton-antiproton annihilation exceeds, by a factor of almost 2000, the energy involved in an electron-antielectron collision, the annihilation process proceeds much more violently, resulting in a "star" formed by many ejected particles.

The proof of the existence of negative protons represents an excellent example of an experimental verification of a theoretical prediction concerning properties of matter, even though at the time of its proposal the theory may have seemed quite unbelievable. It was followed in the fall of 1956 by the discovery of *antineutrons*, i.e., the particles that stand in the same relation to ordinary neutrons as negative protons do to positive ones. Since in this case the electric charge is absent, the difference between neutrons and antineutrons can be determined only on the basis of their mutual annihilation ability.

30-4 The Elusive Neutrino

In addition to protons and neutrons, which form the nuclei of atoms, and electrons, which form their outer envelopes, physicists have discovered an entire array of other particles which, although not permanent constituent parts of the nuclei, nonetheless play an important role in their properties.

Early studies of radioactive β decay led to the conclusion that there was something wrong with the energy balance involved. While α particles emitted by a given radioactive element always carry a well-defined amount of energy characteristic of that element, β particles from one radioactive element show a wide energy spread, ranging from almost zero up to high-energy values. Since the total energy liberation in the transformation of one atomic nucleus into another is expected to be the same for all nuclei of a given kind, it was suspected that there must be another particle coming out of the nucleus, along with the electron, that carried away the missing balance of energy. Since there was nothing

wrong with the balance of electric charge in β decay, the hypothetical particle had to be electrically neutral—and other considerations made it certain from the beginning that its mass must be very much smaller than the mass of an electron. (Physicists are now sure that it has a zero rest-mass, similar to the photon.) Fermi gave it the name *neutrino* ("little neutral" in Italian). Its lack of charge and mass makes it possible for a neutrino to penetrate matter with the greatest of ease, and a heavy concrete wall is no more effective in stopping a beam of neutrinos than a chicken fence is in stopping a swarm of mosquitos. In fact, it would require a lead shield several light-years thick to reduce a neutrino beam to half its original intensity! Thus neutrinos produced in various nuclear transformations escaped unobserved with their loads of momentum and energy, frustrating physicists and causing discrepancies in the balance of the records of incoming and outgoing energy. But whenever there is a suspicion of a new unknown particle, physicists are as good as Canadian Mounties in getting their man, and the nets were gradually drawn close around the elusive neutrino.

The first experimental evidence of the existence of neutrinos, which were originally introduced as purely hypothetical particles, was provided by the observation of the recoil of the nuclei from which the neutrinos were emitted. The unstable isotope of beryllium, Be^7, which can be produced artifically by means of nuclear bombardment, often emits a positive electron and is transformed into the stable isotope of lithium, Li^7, according to the following equation, in which ν represents a neutrino:

$$_4Be^7 \longrightarrow {}_3Li^7 + e^+ + \nu.$$

Instead of this decay, however, another reaction can occur, in which the Be^7 nucleus is transformed into a Li^7 nucleus by capturing one of the electrons from the inner (K) electron shell of the atom:

$$_4Be^7 + e^-_{\text{atomic}} \longrightarrow {}_3Li^7 + \nu.$$

Indeed, the addition of a negative charge to a nucleus is equivalent to the loss of a positive charge. Since the captured negative electron belongs to the original unstable atom, all that happens here is the emission of a neutrino and, through the conservation of momentum, the recoil of the atom from which it came. Since the neutrino does not produce any visible track in the cloud chamber, it looks as if the Be^7 atom started suddenly to move by itself without any agent responsible for the move. This phenomenon was actually observed in a cloud chamber containing unstable Be^7 atoms, and gave the first supporting evidence for the existence of neutrinos.

But the real test of the neutrino hypothesis came in the attempt to stop the escaping neutrinos in their tracks. And, in spite of the almost

incredible ability of neutrinos to make their getaway, physicists managed in 1955 to stop a few of them, thus finding unquestionable proof of their existence. F. Reines and C. Cowan of the Los Alamos Scientific Laboratory used for this purpose the collision process between neutrinos and protons, in which the neutrino is expended to produce a positive electron and to transform the proton into a neutron:

$$\mathrm{p} + \text{neutrino} \longrightarrow \mathrm{n} + \mathrm{e}^{+}.$$

These two scientists built a giant particle counter that registers neutrons as well as electrons and placed it near one of the nuclear reactors at the Savannah River Nuclear Energy Project. The nuclear reactions taking place in the operating reactor produce a tremendous number of neutrinos that stream out through a heavy shielding which holds back all other nuclear radiations. Although the chance of a neutrino hitting a proton and producing the reaction mentioned above is only 1 out of 10^{30}, some of these reactions do actually take place, resulting in the simultaneous appearance of a neutron and the accompanying positive electron. Thus the uncatchable neutrino was finally caught and joined the company of well-established elementary particles.

30-5 Exchange Forces and Mesons

The next member to enter the growing family of auxiliary nuclear particles was also born as the result of purely theoretical considerations. In 1935, a Japanese theoretical physicist, Hidekei Yukawa (known as "Headache" Yukawa to students who struggle with his mathematics), proposed a new particle which would account for the strong forces binding neutrons and protons together in the nucleus. The idea that a force between two particles could be explained by the introduction of a third particle does not seem to make much sense at first glance. Although the mathematical ideas underlying this idea are quite formidable, a crude analogy will help emphasize the basic simplicity of the concept. If you and a friend stand a few feet apart and throw a heavy medicine ball rapidly back and forth, the total effect will be the same as though there were a force of repulsion between you. If each of you stood on a wheeled platform, you would be quickly pushed apart; as the distance between you increased, the time intervals between throwing and catching the ball would likewise increase. This would mean that the rate of change of momentum, and thus the force between you, would become smaller as the distance grew larger.

A force of attraction is not quite so clearly pictured. However, if you imagine yourself grabbing the medicine ball out of your friend's hands, and him then grabbing it from yours, and so on, the result will be a force of attraction between you. The repellent and attractive forces described above could be termed ***exchange forces***, since they arise from the exchange of the medicine ball.

Yukawa's new particle took the part of the medicine ball in explaining the short-range forces of attraction between nucleons. According to his theoretical considerations, the new particles must have a mass intermediate between that of protons and that of electrons, so they received the name *mesons* (from the Greek *mesos* meaning "between").

Two years after the introduction of these purely hypothetical particles as an explanation of nuclear forces, mesons were actually observed in cosmic-ray studies by the American physicist Carl Anderson. The primary cosmic-ray particles bombarding our atmosphere have energies ranging from comparatively low values to as high as 10^{20} ev. Colliding with the nuclei of atmospheric oxygen and nitrogen at the outer fringes of the atmosphere, these primary cosmic-ray particles produce various kinds of penetrating radiations, including high-energy γ quanta and streams of negative and positive electrons; in fact, as was mentioned earlier, positive electrons were first discovered in cosmic rays. Observing the tracks formed by cosmic-ray particles in a cloud chamber placed between the poles of a strong magnet, Anderson noticed that the trajectories of some of the particles, both positively and negatively charged, were bent by a magnetic field more than would be expected in the case of fast protons but considerably less than should be the case with electrons. From the observed magnetic deflection, Anderson estimated that this new kind of particle was about 200 times heavier than an electron, which was in agreement with Yukawa's theoretical prediction. The behavior of the new particles, however, in their reluctance to react with nucleons, made it very doubtful that these were the predicted exchange-force particles. Ten years later, the British physicist C. F. Powell demonstrated that there were two kinds of mesons: the π meson (*pi meson*, or *pion*), which is produced at the upper fringes of the atmosphere by primary cosmic rays, and the μ meson (*mu meson*, or *muon*), into which the pion spontaneously decays in about 10^{-8} sec after being formed:

$$\pi^{\pm} \longrightarrow \mu^{\pm} + \text{neutrino}.$$

In addition, there are also neutral pions (π^0). Neutral pions possess a very short lifetime (about 10^{-16} sec) and, in spite of their high velocity, break up into two γ quanta,

$$\pi^0 \longrightarrow \gamma + \gamma$$

long before reaching the surface of the earth. Positive, negative, and neutral pions interact very strongly with atomic nuclei and are apparently the particles introduced hypothetically by Yukawa for explanation of nuclear forces.

For the study of pions and their decay, photographic equipment attached to large balloons was sent high into the stratosphere. Since cloud chamber equipment is too bulky and heavy to be sent up in balloons, cosmic-ray researchers developed a new method for photographing the

tracks of cosmic particles at high altitudes. Instead of using the ionizing properties of fast-moving charged particles passing through humid air, the new method is based on the fact that these particles affect the grains through which they pass when they travel through a very thick fine-grained photographic emulsion. When the photographic plate is developed, it shows dark streaks that correspond to the trajectories followed by the particles. A very rare photograph of this kind showing the formation of a pion resulting from the collision of a primary cosmic-ray particle with a composite nucleus and the subsequent decay of this pion into a muon and an electron is shown in Fig. 30-4. After the development of the Bevatron and the Cosmatron, it became possible to produce pions in the laboratory, thereby considerably accelerating the progress of the study of their properties.

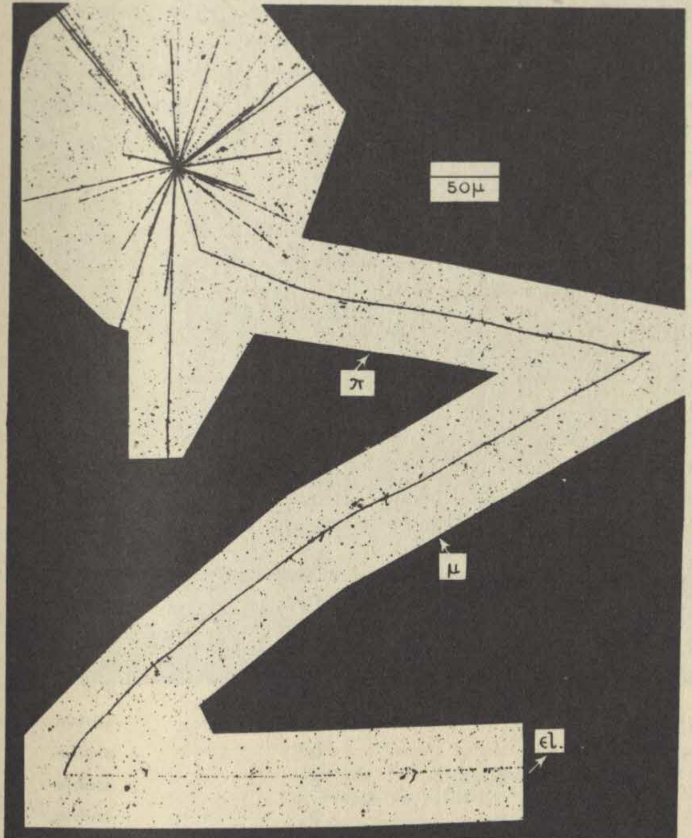
Muons (which are now known not to be mesons at all, but merely an inexplicable heavy edition of the electron) have been detected by cloud chambers in relatively large numbers at the surface of the earth, and investigations have shown them to have a half-life of about 10^{-6} sec. They decay into an electron and two neutrinos, according to the equation

$$\mu^{\pm} \longrightarrow e^{\pm} + 2\nu.$$

(In order to avoid being carried into too many complications, the above equation has been somewhat simplified. Actually, the two neutrinos are different. It was discovered in 1963 that there are two kinds of neutrino: the first associated with the electrons of β decay, the other associated with particle decays involving muons. And each of these has a corresponding antineutrino.)

Muons created high in the atmosphere by the impact of cosmic-ray particles are able to reach the earth's surface in large numbers in spite of their short half-lives. According to relativity theory, 3×10^{10} cm/sec, or 186,000 mi/sec, is the absolute speed limit for any particle. Even at this speed of light, the muon would require 10^{-3} sec to reach the earth's surface if we assume it to have been created at an altitude of 186 miles. If the mean lifetime of a muon is only 10^{-6} sec, how does it manage to survive the journey?

The answer to this question comes from the same relativity theory that enforces the speed limit. You will recall that from our point of view, the watch of a swiftly moving observer will run slowly; relative to us, the muon is traveling at enormously high speed, and the "watch" that times its life appears to us to be running slowly. In fact, the cosmic-ray muons travel at a speed so near the speed of light that time is expanded by a factor of many thousands, and what we measure as a thousandth of a second will be to the muon much less than its lifetime of



Photograph by Dr. E. Pickup, National Research Council, Ottawa, Canada.

FIG. 30–4 A series of nuclear events in a photographic emulsion. A cosmic-ray proton strikes an atomic nucleus (upper left) and produces a burst of many different fragments. One of the fragments, a pion, travels to the right edge, where it decays into a muon and a neutrino. (Neutrinos leave no tracks in cloud chambers, emulsions, or any other similar devices.) In the lower left corner the muon decays into two neutrinos and an electron which then travels to the right.

a millionth of a second. From the point of view of the muon, his watch is running quite normally, but the distance from upper air to surface (which we measured to be 186 miles) may be only a hundred yards or so, which gives it plenty of time to make the trip. Besides explaining the survival of the muon, this argument provides another strong proof for Einstein's Special Theory of Relativity.

30-6 More and More Particles

After the discovery of pions and muons, other particles began to turn up. They appeared in cosmic-ray studies and in experiments with new high-energy particle accelerators. Some of these particles (the *K mesons*) are intermediate in mass between electrons and nucleons, while others (Λ, Σ, Ξ, and Ω particles) are more massive than nucleons and are

known as *hyperons*. In Fig. 30-5, a bubble chamber photograph shows the production and decay of several of these particles.

Table 30-1 lists the conventional "elementary" particles as they are now known. Counting all the antiparticles, which are not specifically shown, there are 34 of them.

TABLE 30-1 THE PROPERTIES OF THE ELEMENTARY PARTICLES OF MATTER (boldface type indicates the particles known before 1930)

Name and Symbol	Mass (in Electron Masses)	Mean Lifetime (in Seconds)
Omega $\Omega^{\pm}$	3284	10^{-10}
Xi $\Xi^{\pm}$, Ξ^{0}	2585	10^{-10}
Sigma $\Sigma^{\pm}$, Σ^{0}	2330	10^{-10}
Lambda Λ^{0}	2182	2.7×10^{-10}
Neutron n	1838.6	10^{3}
Proton $p^{\pm}$	1836.1	stable
K meson $K^{\pm}$, K^{0}	966.5	10^{-8}
Pion $\pi^{\pm}$	273.2	2.6×10^{-8}
Pion π^{0}	264.2	10^{-16}
Muon $\mu^{\pm}$	206.7	2.2×10^{-6}
Electron $e^{\pm}$	1	stable
Neutrino ν_e	0	stable
Neutrino ν_μ	0	stable
Photon γ	0	stable

Also not shown are many of the so-called *resonances*, massive particles with very short half-lives—as short as 10^{-23} sec. If these were to be included, the list of "elementary" particles would total nearly 200. It is becoming more and more questionable to consider all of this growing list as "elementary particles." Many physicists believe that this complex array must be merely different aspects or different energy states of something simpler. So far, however, the needed clues to the mystery have eluded nuclear physicists.

There is some theoretical reason to suspect that "really" elementary particles called *quarks* may exist. These are thought to have fractional amounts of the basic electron charge. If these exist, they must be very seldom alone; they would ordinarily combine to form other particles. Whether they really exist is not known, but they are being looked for in cosmic-ray records. The whole problem of elementary particles is on the frontier of today's physics, and its solution will lead to a much deeper understanding of the nature of the universe we live in.

FIG. 30–5 Most of the bubble trails in this bubble chamber photograph (left) are caused by particles merely passing through and by torn-off electrons. The photograph, however, also shows one unusual sequence of events. As indicated by the diagram on the right, which shows the relevant tracks in the photograph, a π^- meson collides with a proton in the chamber liquid at point A. These two particles, plus the energy of the meson, produce a Σ^0 and a K^0. Almost immediately the Σ^0 decays into a Λ and an electron pair. These two decays have taken place so rapidly that the electron pair shown appears to be coming from point A. The K^0, which produces no ionization, decays at point B into a pair of π mesons. The Λ^0, likewise leaving no trail, decays at point C into a proton and a π^-. The diagram below shows the decay scheme.

$$\pi^- + p^+ \xrightarrow{(A)} \Sigma^0 + K^0$$

$$\Sigma^0 \xrightarrow{(A)} \Lambda + e^+ + e^-$$

$$\Lambda \xrightarrow{(C)} \pi^- + p^+$$

$$K^0 \xrightarrow{(B)} \pi^- + \pi^+$$

The entire chamber is in a strong magnetic field perpendicular to the plane of the picture. Notice that this causes positively charged particles to curve to the right, and negatively charged particles to curve to the left. The amount of curvature depends on the speed of the particle, as well as its mass. The photograph was taken by the Lawrence Radiation Laboratory of the University of California, and is reproduced with permission.

Questions

(30-2) 1. A cosmic ray particle from space with a kinetic energy of 10^{15} ev strikes a nucleus in the upper atmosphere. Assuming that all its energy ultimately goes into electron production, how many electron-positron pairs can be produced?

2. A pattern of detectors scattered over several acres all simultaneously register a sudden burst of electrons of such intensity that it is estimated about 3×10^{10} electrons and positrons were in the shower. Assuming that this entire shower was caused by one cosmic ray particle, what is the least possible kinetic energy it could have had?

(30-3) 3. In bombarding a target with a 6.2-Bev particle to make a proton-antiproton pair, how much energy is left over after the collision, representing KE of particles, photons, etc.?

4. When a proton and an antiproton annihilate, what is the wavelength of the photons produced?

(30-5) 5. Consider a muon traveling at a speed of $0.999c$. How far (on the average) will it be able to travel before decaying into an electron and two neutrinos?

6. What distance would a π^0 pion to be able to travel (from our point of view), if it had a speed of $0.9999c$?

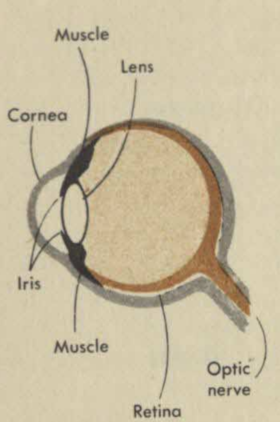
chapter / thirty-one

Biophysics

31-1 Accommodation of the Eye

Lenses, telescopes, and microscopes are man-made devices designed to help the human eye see very distant or very small objects. But the eye itself is an optical instrument designed by Mother Nature, the most ingenious of all designers, and it still remains the most versatile optical instrument ever made. The eye (Fig. 31-1) is essentially a double-lens system, the first and more important lens being formed by the *cornea* and the second by a deformable, transparent cartridge usually known as the *lens*, although "auxiliary lens" would be a more suitable name. In the eye disease known as a "cataract," the lens becomes opaque, but vision can be restored by removing the lens and compensating for its loss by a glass spectacle lens outside the eye or a plastic contact lens on the surface of the cornea.

By the combined action of the cornea and the lens, the image of the object at which we look is formed on the back wall of the eyeball, which is covered with a multitude of nerve endings sensitive to light and which is known as the *retina*. Muscles supporting the lens can increase its



curvature at will so that the eye can be focused on objects located at different distances. There is no ocular musculature to make the lens *less* convex; accommodation appears to depend entirely on squeezing the lens to make it stronger. The normal eye, when perfectly relaxed, is accommodated for infinity; i.e., objects a long distance away are clearly focused on the retina. To focus on nearby objects, it is necessary to contract the muscles of the eye lens in order to make the focal length of the lens shorter. The lens of the farsighted eye is not sufficiently strong to focus even distant objects when it is relaxed; thus a farsighted eye must exert a constant squeeze on the lens muscle, and consequently tires easily. The lens of a myopic, or nearsighted, eye is too strong, so that distant objects are focused in front of the retina when the eye is relaxed.

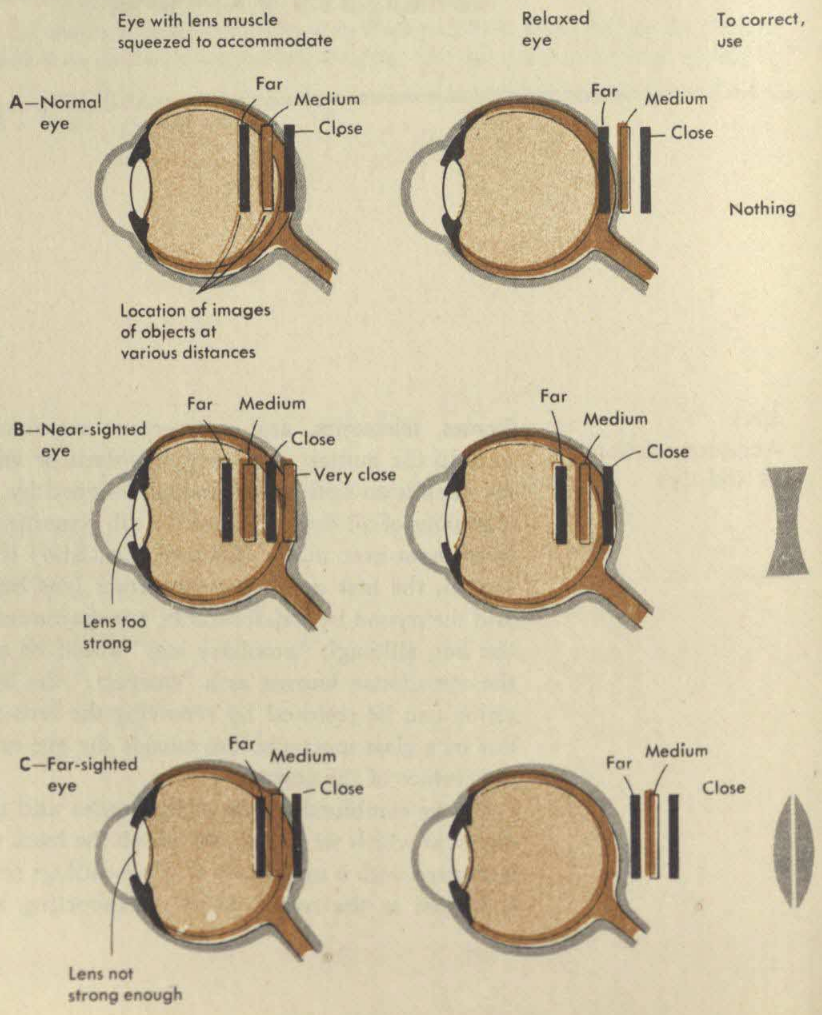


FIG. 31-2 The positions of the image of objects that are at far, medium, or close distances from normal and from abnormal eyes.

It cannot by any muscular effort accommodate for distant objects and hence has a relaxed lens muscle and a blurred image it can do nothing about. In young, normal individuals, the accommodation power of the eye (i.e., its ability to focus at different distances) is very high, and a child can usually accommodate his eyes for clear vision even if the spectacles belonging to a very nearsighted person are put on his nose. With age, the accommodation power of the eye decreases, and its normal focusing abilities very often become defective.

Figure 31-2 shows the focusing of accommodated and relaxed normal eyes, nearsighted eyes, and farsighted eyes. In nearsighted persons, the eye lens is curved too much, so that the image is formed generally in front of the retina. While the eye can still accommodate for good vision of objects located quite close, it cannot possibly bring to the retina the image of distance objects, and they appear blurred. On the other hand, the eye lens of a farsighted person does not have enough curvature, and the image is formed, on the average, behind the retina. In this case, the eye can still accommodate for the clear vision of distant objects, but fails to do so for objects that are comparatively near. Opticians can help in both cases by prescribing concave spectacles for nearsighted persons and convex spectacles for farsighted ones.

Take, for example, a nearsighted person who cannot focus clearly on objects more than 36 in. away. What must be the focal length of glasses that will permit him to focus on very distant objects? The problem is a simple one if we rephrase it to ask: What focal-length lens will produce a virtual image 36 in. to the front, where the eye can use it as a virtual object? To answer this we need only write (since $1/p + 1/q = 1/f$)

$$\frac{1}{\infty} + \frac{1}{-36} = \frac{1}{f}$$

$$f = -36 \text{ in. focal length.}$$

31-2 Color Vision

If you look at a continuous spectrum through a spectroscope, or at the spectrum of sunlight spread out crudely by a simple prism in a darkened room, you will notice that the eye is not equally sensitive to all colors. The spectrum of sunlight represents a quite even distribution of energy from the invisible ultraviolet through the visible and into the infrared. To our eyes, however, the spectrum appears brightest in the yellow-green and fades out into invisibility at the far violet and the far red. Figure 31-3 shows the overall sensitivity of the eye to the various wavelengths included in the visible spectrum.

The light-sensitive nerve endings that thickly cover the retina are of two kinds: the *rods*, which are very sensitive to dim light but have no perception of color; and the *cones*, less sensitive than the rods, but which

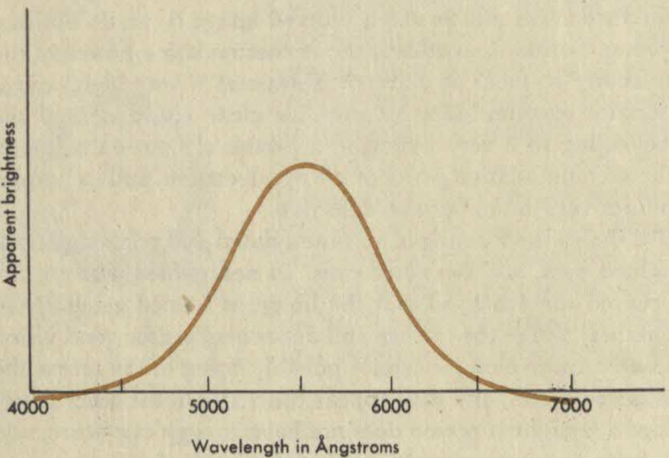


FIG. 31–3 The overall sensitivity of the eye to radiation of various wavelengths.

provide us with our ability to distinguish colors. The exact details of color vision are not yet certain, and the most widely accepted theory is still that of Thomas Young and Hermann von Helmholtz, dating from early in the nineteenth century. There are three types of cones on the retina, each type having a different sensitivity to light of different wavelengths. Each type of cone responds to a wide range of wavelengths (Fig. 31-4), but receives its name from the color to which it is most sensitive. With the possible exception of the deep red beyond 7000 Å, light of any single wavelength will stimulate at least two types of cones, and, for wavelengths shorter than about 5000 Å, all three types are stimulated. Our perception of color seems to depend on the relative strengths of stimulation of our three types of cones. Many different mixtures of wavelengths will provide the same relative stimulations and will thus be judged to be the same color. For example, in Fig. 31-4, a pure monochromatic light of about 5750 Å will stimulate both the green and the red receptors equally (as represented by the distance AB). If light of two different wavelengths, 5100 Å and 6100 Å, and of equal intensity were to strike the eye, the color sensation would be the same as for the 5750-Å light. The 5100-Å light stimulates the green receptors to the extent CE and the red receptors to the extent CD. The 6100-Å light stimulates the green to the extent FG and the red to FH. Graphical addition shows that $CD + FH = CE + FG$; the red and green are equally stimulated, and so the color will appear identical with that of 5750 Å. Thus, as far as the eye is concerned, it does not make any difference whether the incident light is "really yellow" or a mixture of green and red colors.

In spite of the general acceptance of the Young-Helmholtz theory of color vision as described above, there have been many minor technical

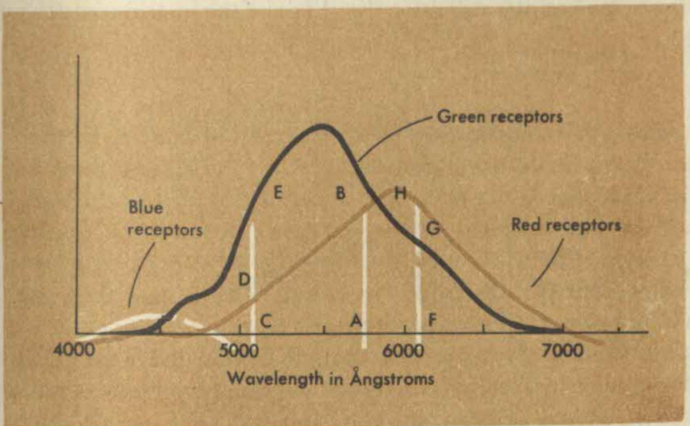


FIG. 31–4 The sensitivity of each of the three types of receptors to radiation of various wavelengths.

objections to it, and its acceptance has not been without reservation. In 1959, Edwin H. Land, inventor of Polaroid and the Polaroid Land Camera, published results of experiments which throw this whole theory into some doubt,* and indicate that our ability to see colors is an even more complicated business than we had supposed. It is apparently possible to reproduce our whole range of color sensations by superimposing only two colors. It is not even critical what these colors are, except that, in general, one be of long wavelength and the other of short; even a combination of red and white light gives excellent color reproduction. The relative proportions of the long and short waves at any point contribute to our sensation of color at that point, but the color we perceive there is influenced by what surrounds the point. Apparently, subjective interpretation by the mind is almost as important to color vision as is the actual combination of wavelengths. Years of further research by physicists, opticians, biochemists, psychologists, and others will no doubt lead to a fuller understanding of our remarkable ability to perceive colors.

Under very dim illumination, such as moonlight, we see little or no color; the whole world consists only of shades of gray. This is because the faint light cannot stimulate the less sensitive cones, and our seeing is being done with the rods, which have no ability to discriminate colors.

Although we may criticize the eye with respect to its ability to distinguish between different incident wavelengths of light, we should not forget that its main function is to inform us about the position and shapes of objects emitting light. It is true that the human ear can do much better in analyzing the complexity of sound emitted by a symphony

* *Scientific American*, **200** (May 1959), 84; and **201** (September 1959), 16.

orchestra, but, on the other hand, we cannot "hear" the location of the individual musicians on the stage nor the shape of the instruments they are playing!

31-3 Speech and Hearing

In human relations, the primary role of sound is, of course, to carry conversation: one person speaks and another listens. These sounds are produced by vibrations in our *vocal cords*, which are located at the opening of the throat and modified by the resonance action of our mouth and nasal cavities. When a person is silent, his vocal cords are wide open, a position that permits the air to circulate freely and facilitates the process of breathing. When a person starts to speak or sing or shout, the vocal cords come close together and begin to vibrate under the action of the stream of air expelled from the lungs. By changing the tension in the vocal cords, we can regulate the pitch of the sounds we produce.

The sounds that come from the human throat during an ordinary conversation are not the pure musical tones that are produced by tuning forks. They are a conglomeration of many different frequencies, and the way these frequencies are mixed together determines whether a sound is "Ah," "Oh," or "Rrr . . . ." These many mixed frequencies arise from the irregularly shaped cavities of the throat and mouth. The cavities, as a result of their different dimensions in different directions, reinforce some frequencies and allow others to die out. By unconsciously regulating the size and shape of our resonant cavities, we control which frequencies are to be favored and which suppressed and thus determine the sort of sound we make.

The importance of the overtones can be easily demonstrated by taking a couple of breaths from a balloon filled with helium* and then speaking in a normal tone. Your voice will sound high-pitched and squeaky, and will be difficult to understand. The fundamental frequency, however, is unchanged; it is controlled by the tension in the vocal cords, which is still the same. And since you are shaping your mouth and throat in the same way to speak familiar words, there is no change here, either. The standing waves in these cavities will have the same wavelength, then, with helium as they had with air. But—what is the velocity of sound in helium? We saw in Chapter 9 that the velocity of sound in a gas depends on the density of the gas; the lower the density, the higher the speed. So in helium, with its low density, the velocity is much greater than in air. Hence, from $v = f\lambda$, with the same wavelengths, the higher velocity means a higher frequency for all the overtones. The voice sounds high-

---

* **NOT** hydrogen! A mixture of hydrogen and oxygen in one's lungs fills them with a violent explosive. The slightest accidental electrostatic spark anywhere in the supply system would almost certainly transfer the experimenter from the tax rolls to the obituary column.

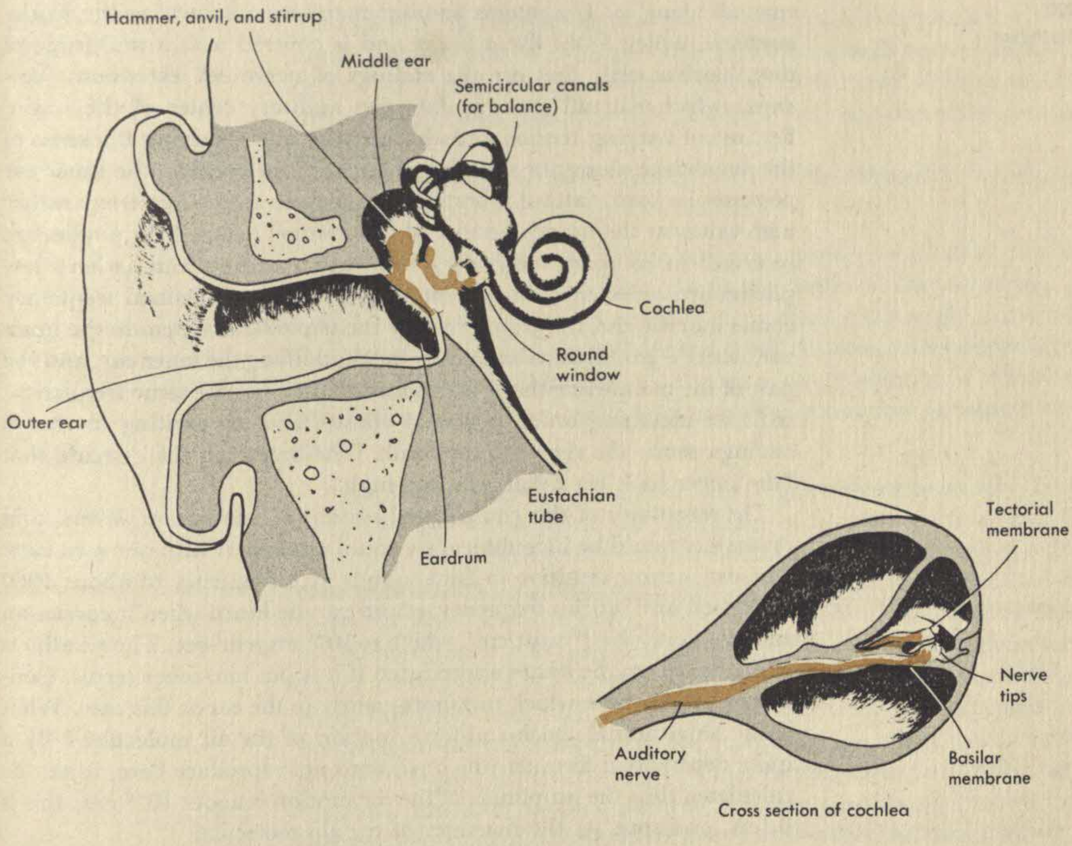


FIG. 31-5 Structure of the human ear.

pitched and unnatural, even with the same fundamental frequency.

The *ear* has a rather complicated structure (Fig. 31-5), consisting of three different compartments: The *outer ear* is a chamber (though not a very effective one) for catching the sound, and at its inner end is the *eardrum*. The *middle ear*, which is connected by a channel with the mouth cavity, contains a joint system of three bones known as the *hammer*, the *anvil*, and the *stirrup*, names that describe the shape of these bones remarkably well. The *inner ear*, which receives vibrations from the eardrum via this triple-bone system, contains the important structure known as the *cochlea*, which looks like a snail shell. In contrast to the outer and middle ears, the inner ear is filled with a fluid that transmits the pressure of the sound waves brought in by the three bones. As shown in the cross-section drawing in Fig. 31-5, the long spiral-shaped tube forming the inner ear contains a membrane known as the *tectorial membrane*, which

runs all along it. It contains another membrane known as the *basilar membrane*, which looks like a brush and is covered with a multitude of tiny, hairlike cells that are the endings of nerve cell extensions (*dendrites*) which run all the way into the auditory center of the brain. Because of varying tension and also because of the varying thickness of the membrane along the spiraling channel of the cochlea, the inner ear possesses its own natural vibration frequencies that range from rather high values at the broad entrance of the channel (where high pitches are received) to comparatively low values at its narrow end (where low pitches are received). When a musical tone of well-defined frequency comes into the ear, it is transferred by the triple-bone system to the inner ear, where it produces vibrations in the fluid filling the inner ear, and the part of the membrane that is in resonance (i.e., has the same frequency) with the incoming sound begins to vibrate and, by exciting the nerve endings, sends the signal to the brain. The brain can then decide that "the singer took his E-flat a bit too high."

The sensitivity of this complicated system of membranes, levers, and resonators would be incredible if we could not hear it with our own ears. The ear is most sensitive to faint sounds at a frequency of about 1000 cycles/sec, and, at this frequency, sound can be heard when it carries an energy of only 10^{-16} watt/cm^2, which is 10^{-9} erg/cm^2-sec. The smallness of the figure can be better appreciated if it is put into other terms. Consider a sound wave which transmits energy to the ear at this rate. What is the longitudinal, back-and-forth motion of the air molecules? By a quite dependable formula which we shall not reproduce here, it can be calculated that the amplitude of the air motion is about 10^{-9} cm; this is 0.1 Å, or about $\frac{1}{40}$ the diameter of an air molecule!

31-4 Biological Effects of Radiations

In connection with the worldwide development of atomic industry and the possibility of atomic warfare, it is very important to know the degree of harm that penetrating high-energy radiations can cause to a human organism. When a fast-moving nuclear particle or a high-energy γ-ray quantum passes through a material body, it produces a certain amount of ionization by knocking electrons from the atoms that it encounters on its way. If the ionized atoms belong to some complex organic molecule, the molecule might be broken up, and even a small percentage of such breakups could lead to the disruption of the proper functioning of the biological system and result in the death of the organism. Thus the biological effects of penetrating radiations can be measured by the degree of ionization they produce when passing through matter. An accepted unit is known as the *roentgen*, which is defined as the dose of radiation that produces 1.6×10^{12} pairs of ions in 1 gram of

atmospheric air. Remembering that 1 gram of air contains 4×10^{22} atoms, we find that the dose of 1 roentgen corresponds to the ionization of 4×10^{-9} percent of all available atoms. In the case of living tissues, the number of ions produced by 1 roentgen is of the same order of magnitude.

In considering radiation damage to a living organism, we should distinguish between two types of damage:

1. *Pathological damage* to the organism exposed to radiation, which, being sufficiently severe, could lead to the death of that organism.

2. *Genetical damage* to the reproductive organs that might not affect the exposed organisms themselves but could do harm to successive generations. A comparatively weak but continuous irradiation of the entire population could even lead to its complete extinction at some future date.

Pathological damage to individuals requires comparatively large radiation doses and, with the exception of atomic warfare, can be expected only in the case of serious accidents in atomic industry. A lethal dose of radiation is considered to be about 800 roentgens, which, when delivered over the entire body, leads to inevitable death. Smaller doses of 100 roentgens or less may produce delayed but still lethal effects such as the development of leukemia and cancer. It is usually assumed that small radiation doses such as 0.1 roentgen per week (the accepted tolerance limit in atomic industry) are quite harmless, since living matter is continuously regenerated at a rate that compensates for the damaged material. This may not be quite true, however, since statistics show that the mean life span of radiologists is five years shorter than that of all other members of the medical profession, even though they take all possible measures to minimize the amount of radiation received by their bodies.

Of much more importance are the effects of radiation on the reproductive organs, since these effects are definitely cumulative, and 100 roentgens distributed over a period of many years are just as harmful as the same dose delivered all at once in one minute. Passing through the genetic cells, nuclear radiations (as well as ordinary X rays) affect the chromosomes of the cellular nuclei and cause changes that may manifest themselves as *mutations* in the progeny. Mutations, except for perhaps one in many millions, are admittedly harmful, and, since they are passed down from generation to generation by the hereditary mechanism, they will lead sooner or later to the death of one of the descendants. Thus, while in the process of Darwinian evolution, which was based on the struggle for existence and survival of the fittest, the few beneficial mutations led to a slow improvement of the species, in a balanced human society where the life of each individual is carefully preserved, mutations

are apt to do a great deal of harm. Under normal conditions, we are all subject to natural mutations that are caused to a large extent by the thermal vibrational motion of our molecules and also by the radioactivity of the ground and by cosmic rays. The total amount of natural radiation received by a person from the date of conception till the age of forty (90 percent of the children in the United States are born to parents below forty years of age) is known to be about 4.4 roentgens. This figure is almost doubled if we take into account the additional dose (4 roentgens) received by an average person during his life from dental and medical X rays.

Compared with these natural and medical doses received by an average person in the course of a lifetime, the "atomic age supplement," caused at the present time exclusively by bomb tests, is rather small. The radioactive fallout over the continental United States in the course of the last few years accounts for less than 0.01 roentgen per person per year, which would give a cumulative effect of only 0.4 roentgen over a period of 40 years. This is equivalent to a 10 percent increase in cosmic-ray intensity and has a much smaller effect than is caused by the difference in elevation between Denver and New York. Now that the United States and Russia are no longer making bomb tests in the atmosphere even this small danger is steadily decreasing.

31-5 Medical Use of Radiations

The medical profession has always been quick to seize on discoveries in the physical sciences and turn them to the personal advantage of mankind. Within months of the discovery of X rays, physicians were utilizing them in examination and diagnosis. The use of X-ray photographs to show internal bone structure is now so commonplace that they scarcely require mention. Bones absorb the X rays much more than do the soft skin and muscle tissue, and thus throw a shadow of their outline on the film. Different kinds of soft tissues absorb X rays differently, and X-ray pictures showing some soft-tissue detail can be made, if a less penetrating longer wavelength radiation is used.

The ability of a material to absorb X rays depends very strongly on the atomic number of the material, the absorption being greater for higher atomic numbers. For example, muscle tissue is composed mostly of oxygen, carbon, and hydrogen, with smaller amounts of nitrogen and other elements. All these elements have atomic numbers that are low, and X rays and γ rays penetrate muscle and other soft tissues with comparatively small absorption. Bones, however, although they contain appreciable amounts of water, are composed largely of calcium phosphate, $Ca_3(PO_4)_2$. Both Ca and P have higher atomic numbers (20 and 15), and for this reason bone absorbs X rays much more strongly than soft tissue does.

In order to see, by X rays, the outline or the functioning of the stomach or intestines, the doctor customarily induces his patient to drink a concoction containing barium sulfate ($BaSO_4$), which is relatively opaque to X rays because of the high atomic number of barium.

By causing several beams of X rays to converge on a malignant growth, it is possible to kill undesirable tissues without causing fatal damage to surrounding normal material. Similar irradiation of tumors may be secured by introducing thin hollow gold needles filled with radon gas. The radon quickly decays through several steps to ${}_{83}Bi^{214}$, which emits high-energy (1.8 Mev) γ rays (called "hard" radiation), the effect of which in ionizing tissue is similar to that of X rays.

31-6 Radioactive Isotopes as "Tracers"

Although large amounts of radiation are invariably harmful to any living organism, smaller amounts, if carefully controlled and administered, will do no perceptible damage. This fact has made it possible for biophysicists and biochemists to deliberately use certain radioactive isotopes in many kinds of biological research.

The radioactive isotope ${}_{53}I^{131}$ (half-life 8 days), for example, is absorbed more strongly by the thyroid gland than by other tissues. Because I^{131} decays by β emission, shortly after a patient has received a dose of it sensitive detectors of the emitted electrons can accurately show the size and location of the hidden gland. Continuing observations of the radiation can supply important information about the thyroid's activity and functioning.

This same isotope of iodine (as well as a few others) has also proved useful in the diagnosis of otherwise almost inaccessible brain tumors. It is incorporated into the complex molecules of certain organic dyes which are absorbed much more strongly by the tumor tissues than by normal tissue. Observation of the emitted radiation can then give the radiologist a good idea of the size and location of the tumor.

Besides these and many other biological applications, tracers find many industrial uses. As an example, a small amount of a radioactive isotope of iron may be incorporated in the steel of a bearing to be tested. After a period of running, the lubricating oil can be checked for radioactivity caused by a microscopic wearing away of the steel bearing. Tests of this sort are very delicate and accurate.

31-7 Thermodynamics of Life

The first law of thermodynamics is essentially a statement of the law of conservation of energy, and living organisms are as much subject to its requirements as calorimeters are. If any object—whether a worm, a man, a heated building in the winter, or the sun—is to maintain a constant temperature, as much heat energy must be lost through its surface as is produced within its volume.

Consider two hypothetical hot objects of the same volume, but of very different shapes; one of them (object A) has, say, a surface area of 100 cm^2, and object B one of 1000 cm^2. If we put both of these objects out on a winter day, B will lose heat (through radiation, conduction, and convection) at ten times the rate of A, because its surface is ten times greater. Thus, to maintain a certain temperature, B will have to generate heat at a rate ten times faster than A.

The same argument can be applied to two objects of the same shape, but of different sizes. Let us put out on the same winter day two cubes; one of them (A) measuring 10 cm on a side, and the other (B) 20 cm. A little arithmetic will show that B has four times as much surface area as A. (This comes readily from the fact that B's linear dimensions are twice those of A; B's surface area will therefore be $2^2 = 4$ times as great.) Therefore, B must produce heat at four times A's rate in order to maintain its temperature. But B has $2^3 = 8$ times the volume of A. Its rate of heat production ***per cubic centimeter of volume*** will therefore be only half that required of A.

We thus see that the surface-to-volume ratio is the criterion for the necessary heat production per unit volume. For bodies of roughly the same shape, this ratio is greater for small bodies than for large. This idea can be readily applied to members of the animal kingdom, all of whom maintain about the same temperature and have about the same density. Hummingbirds, which have a large surface-to-volume ratio, have to metabolize at a terrific rate (a rate, incidentally, that is just about the same as the power production per unit weight in a helicopter). On the other hand, large animals can be very economical in their internal heating system. If an elephant metabolized at the same rate as a hummingbird, it would soon be a roasted elephant, since its body temperature would rise to about that of a kitchen oven. Figure 31-6 graphs the metabolic rate for several different animals. On the graph has been added (but not to scale!) the rate of energy production in the sun, which amounts to about 1.7×10^{-4} cal/gm/hr, averaged over its entire interior.

At first sight, it would seem that the second law of thermodynamics fails in the case of living organisms, since the word "organism" itself implies a high degree of order and organization of the molecules forming it. Consider a plant growing from a seed inside a glass box containing soil at the bottom and abundantly supplied by fresh water and air, which brings in the carbon dioxide necessary for building new material for the growing plant. Although both water and carbon dioxide gas possess a very low degree of order, the plant nevertheless manages to organize H, O, and C atoms and to form these simple substances into highly complex organic compounds such as sugars, proteins, etc. Don't

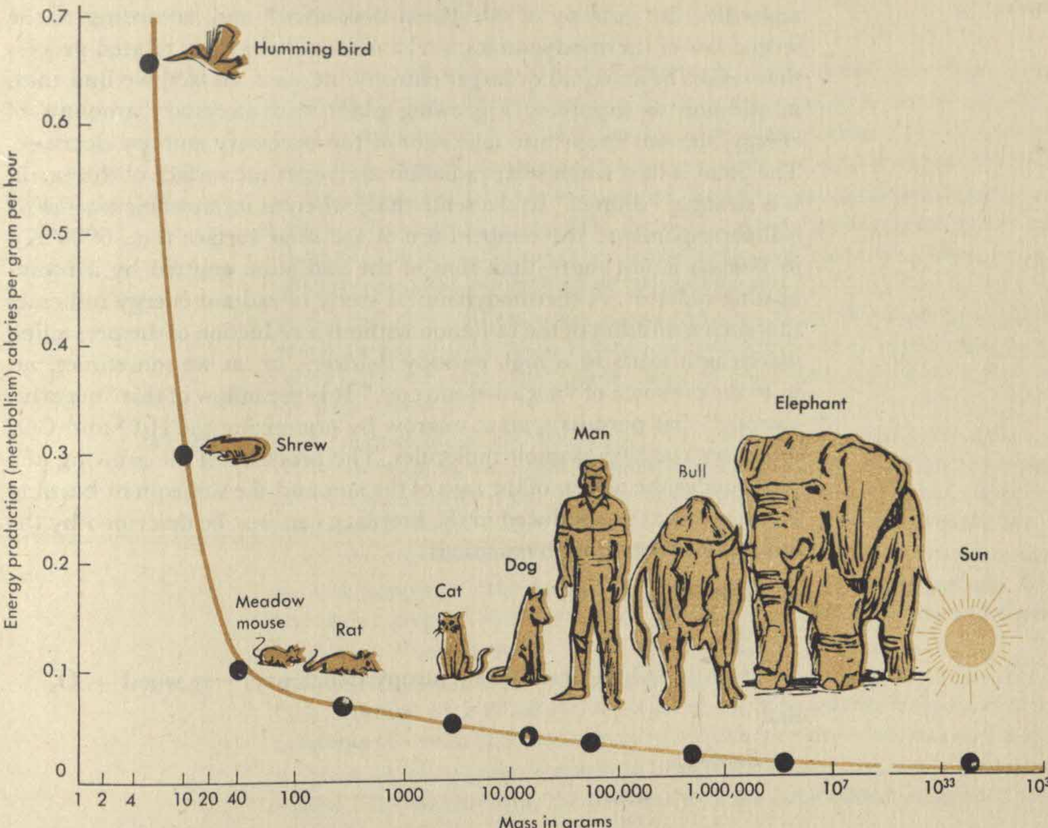


FIG. 31–6 The relationship between mass and rate of heat production in animals ranging in size from hummingbird to elephant. The graph is extended far to the right to include the sun.

we witness here a case in which the degree of order or organization spontaneously increases, the entropy decreasing in contradiction to the basic rules of thermodynamics? The answer is: No. We have forgotten that, apart from water and a carbon dioxide supply (plus a small amount of some salts from the soil), the plant needs for its growth and development an abundance of sunshine. The sun's rays that are absorbed by the green leaves of the plant bring in the energy necessary for building up complex organic molecules from the simple molecules of H_2O and CO_2; but this is not the point. We are not worried here about the first law of thermodynamics (i.e., conservation of energy), but about the second law, which prohibits any spontaneous decrease of entropy.

During the transformation of H_2O and CO_2 into complex organic molecules, the entropy of the system decreases,* and, according to the second law of thermodynamics, we have to look for some related process that results in an equal or larger entropy increase. In fact, we find that, in addition to supplying a growing plant with necessary amounts of energy, the sun's rays also take care of the necessary entropy decrease. The point is that when solar radiation arrives at the surface of the earth, it is strongly "diluted" in the sense that, whereas its *prevailing wavelength* still corresponds to the temperature of the solar surface (i.e., 6000°K), its *intensity* is not more than that of the radiation emitted by a room-heating radiator. A thermodynamical study of radiant energy indicates that such a dilution of the radiation without a reduction of the prevailing wavelength leads to a high entropy *deficiency*, or, as we sometimes put it, to the presence of "negative entropy." It is the inflow of this "negative entropy" that permits a plant to grow by organizing the H_2O and CO_2 into more complex organic molecules. The processes of the growing of a plant under the action of the rays of the sun and the subsequent burning of the material so produced in the fireplace can best be described by the following two symbolic equations:

Growing:

$$\underset{\text{from soil}}{H_2O} + \underset{\text{from air}}{CO_2} + \underset{\text{from the sun}}{(\text{energy} + \text{entropy deficiency})} \longrightarrow \text{wood} + O_2$$

and

Burning:

$$\text{wood} + \underset{\text{from air}}{O_2} \longrightarrow (\text{energy} + \text{entropy excess}) + H_2O + CO_2$$

The energy, which *is to be conserved*, is absorbed in the first process and liberated in the second. The entropy, which *must always increase*, increases in both processes, since the "disappearance of a deficiency" is equivalent to the "appearance of an excess."

The high organization of molecules (negative entropy), obtained by plants from the sun's rays, is then passed, along with the accumulated energy, to animals, which in this case are completely parasitic beings, at least from the point of view of the plants.

* The fact that a piece of wood burns spontaneously when set afire proves that (wood) $\longrightarrow H_2O + CO_2$ is the *natural* direction of the process, in which the entropy *increases*.

Questions

(31-1) **1.** A farsighted man cannot focus clearly on objects closer than 48 in. away. What focal length lens will permit him to read a paper at a distance of 18 in.?

2. A nearsighted man cannot focus clearly on objects farther away than 60 cm. What focal length lens will enable him to see clearly objects at a great distance?

3. An elderly man who has lost nearly all ability to accommodate wears bifocal glasses. The upper part, through which he looks at distant objects, has a focal length of $+50$ cm. What is the focal length of the lower part of his glasses, through which he can read a book held 30 cm from his eyes?

4. A nearsighted person wears glasses of $f = -50$ cm. Through these glasses he has perfect distant vision, but cannot focus clearly on an object closer to his eyes than 25 cm. How close can be bring an object and still see it clearly if he takes his glasses off?

(31-2) **5.** A mixture of red and blue light gives the eye the sensation of purple. (a) Why can there be no single frequency which gives the sensation of purple? (b) Why does a piece of purple cloth look very different in sunlight and under an incandescent lamp? (This difference occurs in all colored materials, but is especially conspicuous with purples.)

6. Two samples of cloth appear equally bright when viewed in sunlight. One sample is a greenish-blue, the other a reddish orange. Which of two will appear brighter in the light of an incandescent lamp?

(31-3) **7.** The velocities of sound in different gases vary inversely as the square root of their densities. Air is 8 times as dense as helium. How will the frequencies of the overtones of a voice in helium compare with the same overtones sounded in air?

8. How will the overtone frequencies of a voice speaking with xenon (molecular weight 131) compare with the frequencies of the same overtones sounded with air? (av. mol. wt. = 29.)

(31-7) **9.** Assume a 10-gm shrew and an 80-kg man each to have a surface area twice that of a sphere of the same volume, and to have a density equal to that of water. (a) What are the surface/volume ratios for these two animals? (b) From Fig. 31-6, how do the rates of energy productions per unit of surface area compare for man and shrew? (c) How does the s/v ratio depend on the radius of the assumed sphere? (d) If coarse rock, measuring about 10 cm in diameter, is crushed into a powder whose particles have an average diameter of 0.1 mm, by what factor has the surface area been increased?

10. We have seen that, in spite of its very high surface temperature, the energy production of the sun per gram is very much lower than that of any living organism. But how does its energy production per unit of surface area compare to that of a shrew or a man? (See Question 9.)

chapter / thirty-two

Geophysics

32-1 The Deeper, the Hotter

It is well known to mining engineers that as they dig deeper and deeper into the earth's crust they encounter a steady rise in temperature. The most comprehensive information concerning the rise of temperature with increasing depth under the surface of the earth is obtained from deep-well boring, which has been carried on in several thousand localities all over the surface of the globe. Measurements made in these wells indicate that *the rate of temperature increase underground amounts to about 30°C for each kilometer of depth.*

Since the temperature of rocks immediately under the surface (such as in caves) is about 20°C, it follows that, at a depth of only 2.5 km, the temperature of rocks may reach the boiling point of water. In certain localities, the conditions are such that the water penetrates through cracks into these hot regions and, being heated above the boiling point, is thrown up again in magnificent geysers, such as those in Yellowstone National Park (Fig. 32-1) and in Iceland. By extrapolating the observed temperature increase to still greater depths, we find that, at a depth of

Haynes Photo.

FIG. 32-1 Giant Geyser, Yellowstone National Park.

about 50 km, the temperature of rocks reaches 1500°C, which is their melting point. However, due to the tremendously high pressure (about 15,000 atmospheres) that exists at this depth, the rocks apparently do not become really fluid, but acquire the property of *plasticity* and become similar to sealing wax (which breaks into fragments like a piece of glass if dropped on the floor but flows like honey—although much slower—if left alone for a sufficiently long period of time).

In regions of the earth where the crust is weaker than in other places, occasional deep cracks are likely to be formed, and hot plastic material may be squeezed into these cracks and flow slowly upward. Coming closer and closer to the surface, this material will enter a region of lower pressure, will become more and more fluid, and will finally erupt from volcanic craters in the form of red-hot lava.

32-2 Earthquakes

Besides outbursts of volcanic activity, which eject many thousands of tons of flaming lava and enough volcanic ash to bury entire cities (the Roman city of Pompeii being the outstanding example), these subterranean disturbances often take the form of vigorous tremors in the earth's crust that are felt to a larger or smaller degree all over the world.

In the year 1775, a violent earthquake all but annihilated the Portuguese capital, Lisbon, and killed 15,000 people; the Messina, Sicily, earthquake of 1908 cost 100,000 lives; and the San Francisco earthquake of 1906 claimed an estimated 452 lives. The Japanese, who live on what amounts to a powder keg, suffer the most from quakes. The 1923 earthquake alone took a toll in Japan of 99,331 killed, 103,733 injured, and 43,476 missing. On the other hand, British earthquake casualties have so far been limited to a single person, who was killed by a falling stone during the London earthquake of 1580.

Destructive as they are, earthquakes are of great help to scientists in their study of the interior of our globe, since earthquake waves originating in some point of the earth's crust propagate to other points on the surface of the globe through its deep interior. As we have seen earlier, there are two kinds of waves that propagate through a continuous medium:

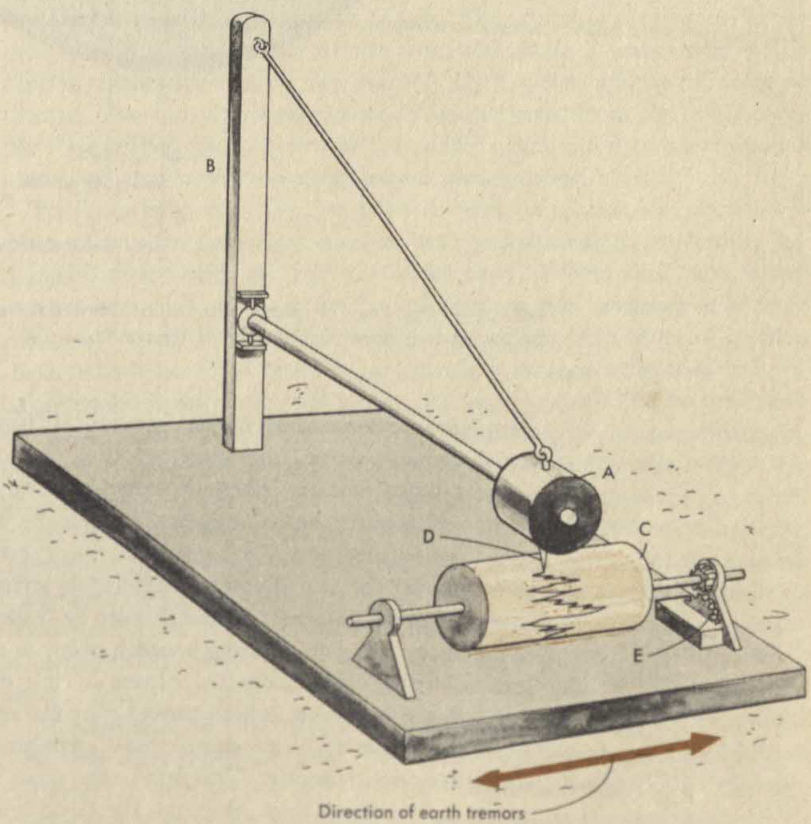
1. *P waves* (pressure or push waves) are longitudinal waves that can propagate equally well through a solid and through a fluid medium.
2. *S waves* (shear or side waves) are transverse waves that can propagate through solids but not through liquids.

When an earthquake wave from a distant disturbance arrives at the surface of the earth, the motion of the ground in the case of P waves will be in the direction of propagation, while in the case of S waves the motion will be perpendicular to it. This permits us to distinguish between

P and *S* waves by registering the movements of the ground by means of a very sensitive instrument known as a *seismograph.*

Seismographs are based on the principle of inertia, according to which each body at rest tends to preserve its state of rest. A simple seismograph, known as the *horizontal pendulum*, is shown in Fig. 32-2 and consists essentially of a heavy weight *A* which can move with very little friction around the vertical support *B*. If the ground on which this instrument is installed is jerked by an earthquake wave in a direction perpendicular to the vertical plane passing through the support and the weight, the weight remains immovable because of its large inertia, and the displacement of the stand relative to the resting weight is registered on the rotating drum *C*. Two such instruments installed at right angles to each other give us

FIG. 32–2 Schematic diagram of a simple seismograph. A massive weight *A* is suspended from a vertical support *B*. C is a rotating cylinder driven by a clock mechanism *E*; *D* is a pen that marks the rotating cylinder as earth vibrations move it back and forth beneath the nearly stationary mass *A*.



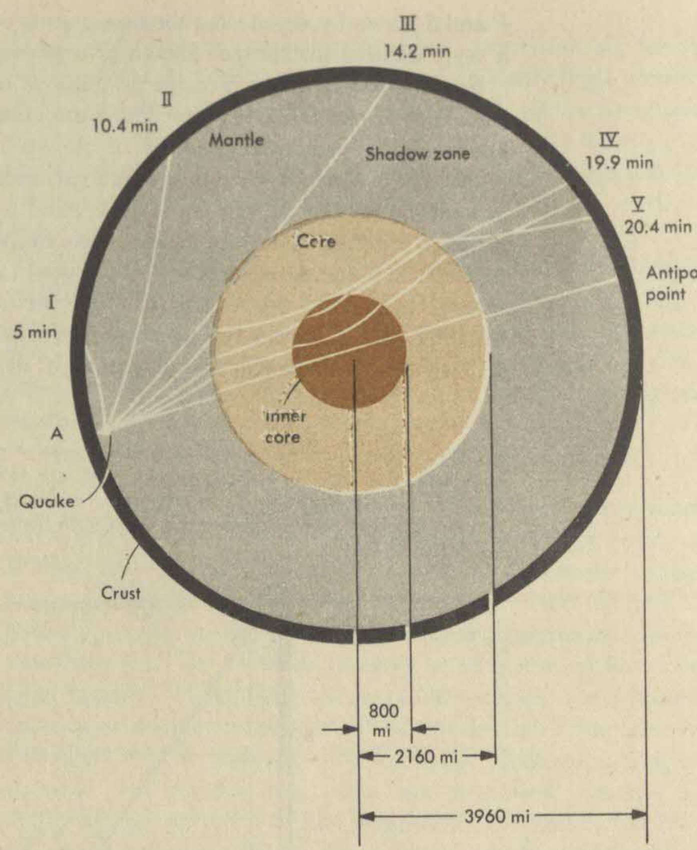


FIG. 32-3 The lines of propagation of an earthquake disturbance originating at a point near the top of the mantle.

complete information about the horizontal displacements produced by the earthquake wave, and there are also seismographs that register the vertical displacements. The seismograms produced by these instruments permit us to carry out a complete analysis of the arriving disturbance.

Let us consider now what would be observed by seismic stations scattered all over the world when a sufficiently strong earthquake originates in some point of the earth's crust. Figure 32-3 shows the propagation lines of the disturbance originating beneath point *A*. At stations I, II, and III, located within 100° from the center of the disturbance, both *P* and *S* waves are observed, which proves that the material of the earth possesses the property of an elastic solid (capable of transmitting a shear) up to very great depths. This fact is the basis for the statements made above to the effect that, although the deep-lying rocks are heated well

above their melting point, they retain the properties of an elastic solid in respect to rapid deformations such as those caused by a propagating S wave.

An amazing point about earthquakes is that *beyond a distance of 100° from the center of the disturbance, there always exists a ring-shaped "shadow zone" withing which the earthquake waves are not felt at all.* Movements of the crust again become noticeable within 35° of the antipodal point (i.e., the point directly opposite the origin of the earthquake), but only those produced by the P waves. To explain this unusual fact, seismologists (i.e., scientists who study earthquakes) had to assume that *the central part of the earth is occupied by a core that is in a liquid state and thus cannot carry S waves.* With this assumption, the existence of the "shadow zone" can be easily understood by inspecting the rays of earthquake waves shown in Fig. 32-3. Due to a rapid increase of pressure and density with depth, the velocity of seismic waves (both P and S) also increases and causes the propagation lines to bend slightly toward the surface. Thus the waves arriving at station III, after passing very close to the surface of the fluid core, are the last ones that still propagate all the way through the plastic material of the rocky mantle. The rays that propagate deeper inward will hit the surface of the core, and while S waves will not be able to propagate beyond this point at all, P waves will be refracted as shown. After the second refraction at the exit point from the liquid core, the P waves will finally arrive at the surface much closer to the antipodal point and thus produce a ring-shaped shadow zone.

The knowledge of the propagation velocities of seismic waves at different depths under the surface enables us to get information concerning the material from which the interior of the earth is formed. It was found in this way that *our globe has an onionlike structure consisting of a large number of concentric shells.* Under a comparatively thin layer of granite rocks, which form the continental massifs, is located a layer of heavier basalt rocks. As was indicated above, this basalt material below the depth of about 30 mi is heated above its regular melting point but remains in a plastic state because of high pressure. Extensive studies of seismic waves propagating through the earth's interior led to the discovery of other layers with interfaces located at depths of 250 and 600 mi. The most interesting discontinuity occurs, however, at the surface of the central core, 1680 mi under our feet, or at 60 per cent of the earth's radius from its center. At this interface, the density, which steadily increases (because of compression) from 3.0 at the surface to about 5.5 at this interface, jumps suddenly to 9.5.

Detailed analysis of earthquake waves passing through the core indicates that there is also an inner core which may, like the mantle, behave as though it were a rigid solid, rather than a fluid like the outer core. The temperature of the core is not known; certainly the increase of 30°C

per mile that we find near the surface does not continue into the deep interior. Most geologists and geophysicists believe the earth's center has a temperature of only a few thousand degrees.

32-3 Why Is the Earth Hot Inside?

Until a few decades ago, it was generally assumed that the earth and all the other planets were formed from the hot material of the sun, and that the earth's interior had just not had time to cool. We know now that this is not so, and that the planets of the solar system were formed by the aggregation of cool interstellar gas and dust particles during the formation of the sun itself. One possible way to explain the heating of the central part of the globe is to ascribe it to the compression of the material under the action of gravitational forces. And, indeed, it can be calculated that the potential energy of gravity liberated in the process of the condensation of a body as large as the earth from a thin dust cloud would be enough to heat its material to a temperature of a few thousand degrees.

But there is another possibility that cannot be overlooked. It is that the heat now contained in the interior of the earth may not be left over from the original formation process of our planet, but quite on the contrary, may be an accumulation that took place after its formation. In fact, we know that the rocks forming the crust of the earth contain a certain amount of natural radioactive elements which liberate nuclear energy at a slow but steady rate. It has been calculated that, if the concentration of these natural radioactive elements in the interior of the earth were about the same as we find on its surface, the amount of heat liberated since the formation of the earth would be large enough to turn it into a flaming red-hot sphere. Since this is not the case, we have to assume that *natural radioactive elements are considerably rarer in the earth's interior*, which is in good agreement with the observed fact that the deep-lying basalt rocks are considerably less radioactive than the surface granite rocks. One ton of ordinary granite rock contains 9 gm of uranium and 20 gm of thorium, while for basalt rocks the respective figures are only 3.5 and 7.7. However, there is no way of setting a lower limit for the abundance of natural radioactive elements in the earth's interior, so the possibility is not excluded that the molten state of the earth's core is the direct result of this radioactive heating. Many scientists are even inclined to believe that, instead of cooling, our planet is still being gradually heated by radioactive decay.

We have just mentioned that natural radioactive elements are encountered almost exclusively in the thin outer crust of the earth. Why is this so? If the earth was formed by an aggregation of dust particles of interstellar material, its original composition should have been uniform from the center to the surface. The accumulation of radioactivity in the

crust, therefore, must be a secondary effect. One possibility is that the parts of the earth's material which contained radioactive elements were heated above the temperature of their surroundings and floated up toward the surface. Another possible explanation of the high radioactivity of the surface crust is that the radioactive elements were acquired by the earth *after* its formation. In fact, it has been recently demonstrated that the giant stellar explosions known as supernovae produce large amounts of radioactive elements which are scattered through the space of the universe by the force of the blast. Thus it may be that, during some phase of its early history, our solar system passed through such a radioactive cloud and became contaminated by its material.

32-4 Upside-down Mountains

The plasticity of the earth's mantle is of paramount importance in determining the surface features of the earth, with its extensive continental massifs, towering mountains, and deep ocean basins. If the earth's crust were smooth and uniform, like a layer of ice on a frozen lake, there would be no dry land, the entire surface of the globe would be covered by an ocean of uniform depth, and the higher forms of life on the earth would be confined to fish. That this is not so is due to the fact that the earth's crust does not have the same thickness and composition everywhere.

When we look at a high mountain range rising thousands of feet above the surrounding plain, we are inclined to consider it merely as a gigantic excrescence of rock piled on the surface of the earth, much like an artificial hill built up by engineers. Such a primitive view, which regards mountains as wholly a surface feature, was common in geology a century ago, and it was only comparatively recently that students of the earth's surface features came to the conclusion that *most of any mountain range is situated under the surface of the earth.* The discovery of these "mountain roots" going very deep under the earth's surface resulted from the study of the gravitational action of a mountain upon two pendulums suspended on opposite sides of it. According to the law of gravity, we would expect the great mass of a mountain to deflect the two pendulums from the "true vertical," as illustrated in Fig. 32-4. Of course, in this instance, the term "true vertical" is defined, not as a plumb line, but as the direction to the center of the earth given by the observation of stars in the two localities. To the great surprise of the scientists who carried out these experiments for the first time, the observed deviation of the plumb line from the true vertical in the vicinity of great mountains did not confirm the theoretical expectations. In the case of Mount Everest, for example, the observed deviations were about three times smaller than should be expected from its giant mass, while the Pyrenees even seemed to repel the pendulum instead of attracting it!

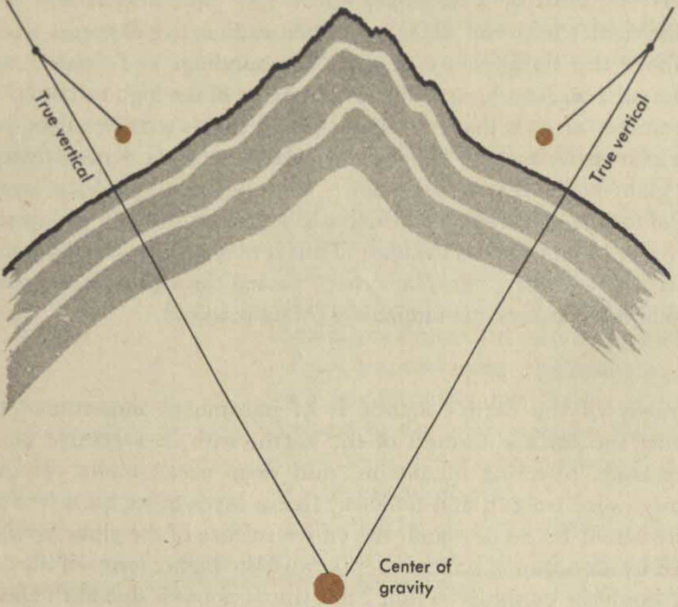


FIG. 32-4 The expected deviation of pendulums from the vertical by the gravitational attraction of a mountain, on the assumption that the earth's crust was of uniform depth and density.

The absence of the expected additional gravitational attraction was originally interpreted as indicating that the mountains were hollow inside, something like giant blisters on the earth's crust. We know now that this interpretation was wrong and that the correct explanation of the observed facts can be traced back to the plasticity of semimolten layers many hundreds of kilometers under the surface of the earth. According to the present view, the mountains on the surface of the earth represent formations similar to the ice hills produced on polar ice floes by the compression of ice. Every polar explorer knows that when crowded blocks of floating ice, broken by compression, are piled on top of one another, most of the ice sinks below the surface to keep the rest afloat. And when a polar bear sees a high hill rising above the surface of an ice floe, a seal swimming under it will notice an even larger bulge protruding down into the water. Similarly, *under each mountain rising above the surface of the earth exists, so to speak, a negative mountain formed by rocky granite material protruding into the underlying plastic layers of more dense semimolten basalt on which it floats* (Fig. 32-5).

According to Archimedes' law, the weight of the floating body must be equal to the weight of the displaced material underneath, so that the presence of an elevation on a floating crust does not signify any actual increase in the weight or mass of material in this region. Accepting this point of view, known as the theory of *isostatic equilibrium*, we can dispose

of the question: "Why does tne mountain not affect the plumb line as much as would be expected from its apparent mass?" We should rather ask: "Why are there any deviations in the plumb line at all?" To answer this question we must bear in mind that the weight of the plumb bob is caused by the attraction between the bob and every separate piece of rock in the whole earth. Obviously, however, the cubic mile of rock directly underfoot will have much more influence than a cubic mile of rock in Australia. If a mountain range rears up above the surrounding plains, it must be floating on an especially thick layer of light granite rock beneath it. This is the material that, being close, has a strong influence on how the plumb line behaves. Since this material is relatively light, its pull on the bob is weak, and the plumb line is deflected less toward the mountain than we would expect.

Although beneath each mountain range there is a roughly corresponding upside-down range of light granite, the strength and elasticity of the earth's crust prevents the upside-down mountains from looking *exactly* like their counterparts above the surface, and an imaginary negative mountain climber making his way through the masses of plastic basalt deep under the Rocky Mountains will look in vain for anything resembling an upside-down Pike's Peak.

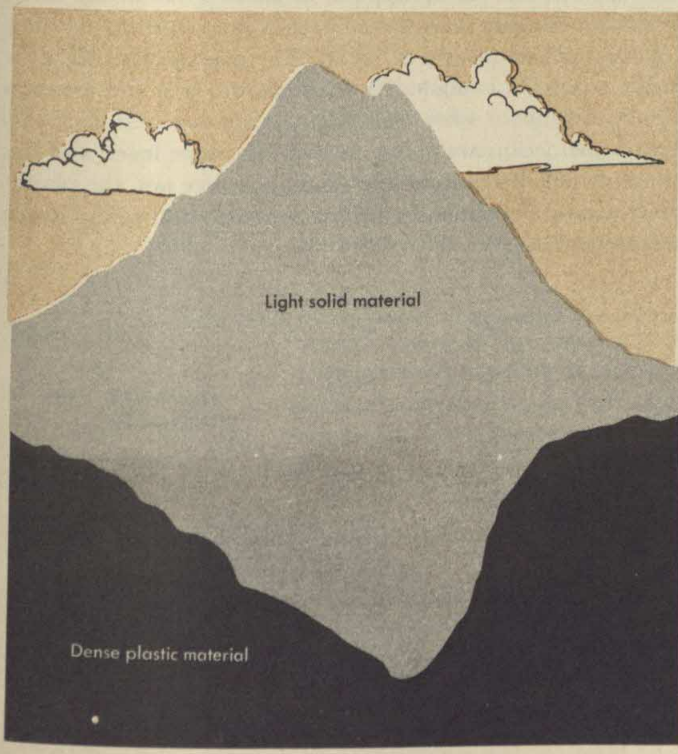


FIG. 32–5 The "upside-down" mountain. The weight of the rocky material protruding above the general level of the earth's surface is supported by the buoyancy on the antimountain beneath it, which floats on the dense plastic material in the interior.

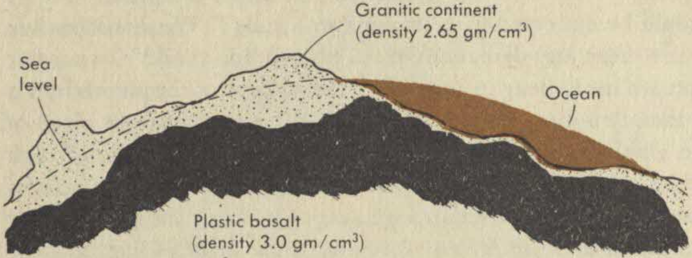


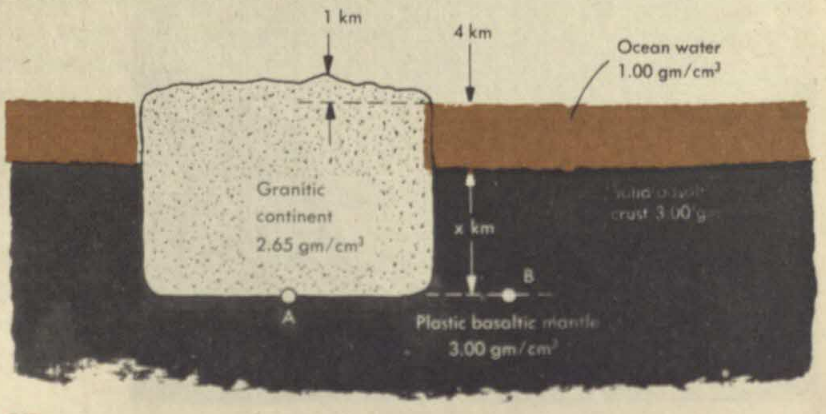
FIG. 32-6 Sketch of the earth's crust, showing a light continent floating on the denser plastic material beneath it.

32-5 Floating Continents

The existence of continents separated by comparatively deep ocean basins also may be ascribed to isostatic equilibrium. Much of the outermost part of the earth's crust is composed of granite with a density of 2.65 gm/cm^3, and the deeper layers of basalt with a density of 3.0. Although in the case of the continental massifs the granite layer is quite thick, this layer is thin or absent at the bottom of the oceans. Thus we may consider continental massifs as solid blocks of lighter material (granite) floating in heavier plastic material (basalt) in much the same way that icebergs float in water (Fig. 32-6).

Archimedes' principle works just as well on a large scale as it does in a laboratory beaker, and can be applied to the structure of the earth itself. Figure 32-7 is a highly schematic diagram of a continent floating on a sea of denser plastic mantle material. The average altitude of a continent above sea level is about 1 km, as shown, and the average depth of the oceans is about 4 km. Point A, beneath the continent, and point B, beneath the ocean, are drawn to be at the same level, and the pressure at these two points must be the same. If it were not, the plastic mantle material would flow from the higher-pressure point to the lower until they were equalized.

FIG. 32-7 Conventionalized diagram of a continent supported by its buoyancy.



So, let us calculate (from $p = hd$) the pressures at A and B and set them equal:

$$p_A = (x + 4 + 1) \times 2.65 = 2.65x + 13.25$$
$$p_B = x \times 3.00 + 4 \times 1.00 = 3.00x + 4.00.$$

So

$$2.65x + 13.25 = 3.00x + 4.00$$
$$0.35x = 9.25$$
$$x = 26 \text{ km}.$$

The thickness of the granite slab forming the continent is then $x + 5 = 31$ km. This is in good general agreement with seismographic evidence. Earthquake waves are partially reflected from the boundary between granite and basalt beneath the continents; by timing the return of these reflected waves, the overlying thickness can be quite accurately determined. Such direct measurements also confirm that the granitic slab is much thicker and deeper beneath mountain ranges such as the Rockies or the Alps than it is beneath the low plains of Kansas or France.

The adjustment of the earth's crust under the shifts of mass on its surface has played a very important role in the evolution of the face of our planet. For example, considerable isostatic adjustment took place during the glacial periods, when thick sheets of ice covered much of North America and Europe. The weight of the ice caused the northern regions of these continents to sink deeper into the plastic layer of basalt underneath. At present, when most of the ice has retreated, the depressed parts of the continents are slowly rising toward their pre-Ice Age level, and we can notice a slow regression of the seas along many shorelines of the northern countries.

32-6 The Rise of Mountains

In studying mountain ranges, geologists long ago noticed that most of them are made up of layers of rock that have been wrinkled up into great folds. Since these layers had originally been deposited as nearly flat sheets of sediment on the bottom of an ancient ocean, it was evident that some enormous force had slowly pushed them up into a series of giant wavy folds. If you push against the opposite edges of a pile of handkerchiefs, you will get a good idea of what must have taken place. This has happened in many places at many times in the earth's history, and it is difficult to escape the conclusion that the earth is shrinking and wrinkling up its crust in the same way that a smooth plum becomes a wrinkled prune as it dries out and becomes smaller.

In the days before the discovery of radioactivity, geologists tried to explain this shrinkage by saying that the earth was cooling. This, how-

ever, was not a very successful explanation. We know that since the temperature of the earth's crust rises about 30°C per kilometer of depth, there *is* a steady flow of heat from the interior out through the earth's surface into space. Physicists know the heat conductivity of basalt and granite, and from this and the known surface area of the earth it can be computed that the earth has lost something like 10^{30} calories in the past 4 billion years. Dividing this number of calories by the mass of the earth (6×10^{27} gm), we find that each gram of the earth on the average has lost about 165 calories. Since the heat capacity of the earth's material is about 0.2 cal/gm/°C, this means that the earth must have cooled by about 800°C since the formation of its solid crust.

How much shrinkage would this amount of cooling cause? If we apply an average figure of 8×10^{-6}/°C for the linear coefficient of the expansion of rocky material, we can, after a little arithmetical juggling, come to the conclusion that it would have shortened the earth's circumference by about 200 mi. This is only enough to wrinkle up one good-sized mountain range, and is not nearly enough to explain the many ranges that have been forming and eroding away during the earth's long history.

Anyway, since the discovery of radioactivity and of the distribution of radioactive elements throughout at least the earth's outer layers, we may expect that the earth's interior is actually getting hotter rather than cooler. Is there any reasonable way in which we can picture the earth warming up on the inside and, at the same time, shrinking? There is, but to see it we must look, not at the surface, but in the deep interior at the boundary between the plastic mantle and the fluid core.

Imagine a planet made of a solid block of ice, with bits of radioactive materials scattered through it. As the radioactive heat collects, the center begins to melt, and the ice planet acquires a core of water that continues to grow as more and more of the mantle melts away at its boundary. Since the density of ice is only about 90 per cent that of water, every cubic mile of ice that melts forms only 0.9 mi^3 of water, and the volume of the ice planet is lessened by 0.1 mi^3. (The great pressure will keep any cavity from forming, of course, and the whole mass of the mantle will slowly shrink inward.)

If this is what is happening to our own earth, we need only substitute for ice some unknown iron compound. The change in density between mantle and core is from 5.5 to 9.5, so that each cubic mile of mantle that changes to the fluid state of the core will shrink in volume by 0.42 mi^3, with an equal reduction in volume for the whole earth.

This may be the answer. Our earth, heated inside by radioactive energy, may be growing smaller as its fluid core grows larger, causing its surface rocks to wrinkle into huge ranges of mountains as it shrinks.

Most geologists, however, now believe that great, slow convection

currents in the plastic mantle material may have provided forces that folded the great mountain chains. Figure 32-8A shows an area of sedimentary rocks overlying the granite crust. At *H* and *H*, heating, due perhaps to some local concentrations of radioactive material, has made the material expand and rise. At *C*, the currents have cooled somewhat, and come together and sink. Many geophysicists believe that the frictional drag between the crust and such slow mantle currents would provide enough force to wrinkle the crust into mountain ranges.

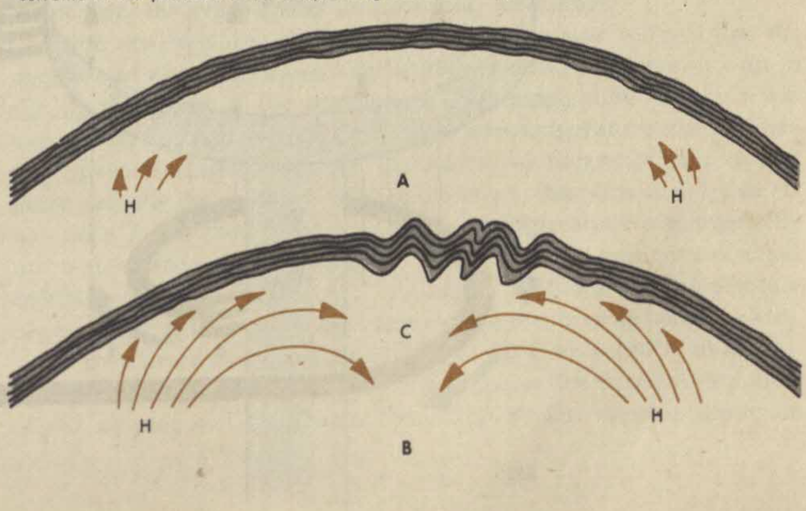
It is humiliating to realize that, knowing as much as we do about the interiors of the sun and other stars, we know so little about the earth's interior, only a few hundred or a few thousand miles beneath our feet!

32-7 Terrestrial Magnetism

The knowledge that pieces of certain naturally occurring iron ores can indicate the direction of the north and south poles existed in China for many centuries and was brought to Europe, along with other oriental rarities, by Marco Polo. Extensive studies of the earth's magnetic field, which is of paramount importance for navigation, led to the construction of magnetic maps from which it became evident that the magnetic poles of the earth do not coincide with its geographic poles. The "north magnetic pole") (which is actually the south pole of the earth's magnet) is located in northern Canada, about 11 degrees 30 minutes from the geographic pole, and the "south magnetic pole" lies at approximately the same distance from the geographic south pole in Antarctica.

The magnetic field of the earth can be described, to a very good approximation, as being produced by a magnet located in its interior and slightly inclined with respect to the rotational axis (Fig. 32-9).

FIG. 32-8 How a mountain range may have been formed by convection currents in the plastic mantle compressing the crustal layers int › folds.



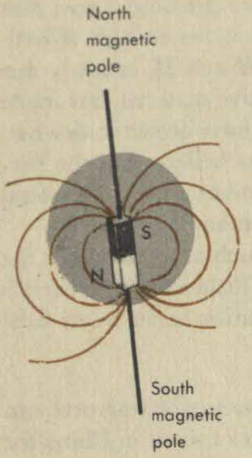
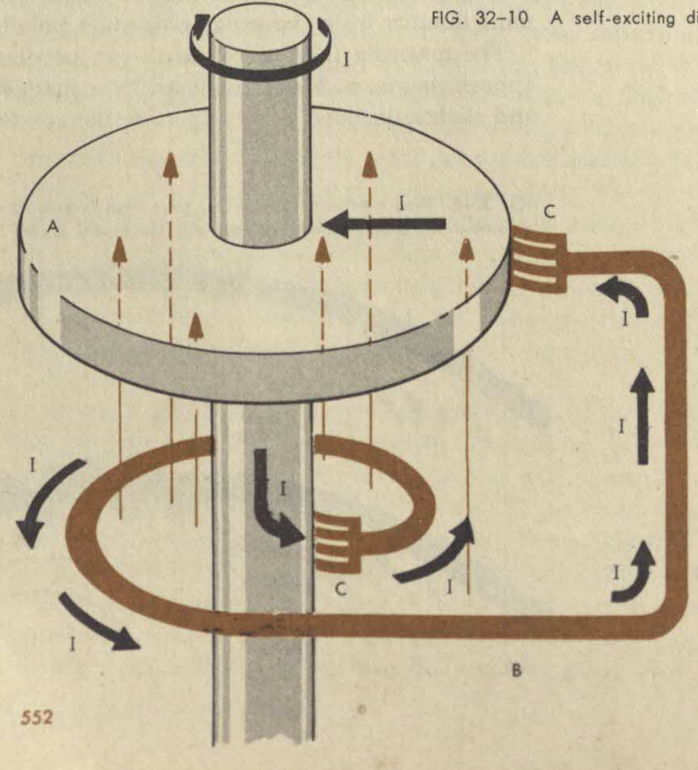


FIG. 32-9 The main features of the earth's magnetic field are similar to those that would be caused by a giant bar magnet (or an equivalent circulating current) at the earth's center.

On this primary geomagnetic field is overlapped a second weaker field of magnetic forces that shows a rather irregular pattern, and which has no apparent correlation with the rotation of the earth around its axis. This secondary field is slowly changing, and its entire pattern is drifting westward at a rate of 0.18° per year. This movement causes slowly varying changes in the direction assumed by magnetic needles on the earth's surface and necessitates the redrawing of navigational maps from year to year. In addition, there is also a magnetic component produced by electric currents in the ionized uppermost layer of the terrestrial atmosphere known as the *ionosphere*.

The origin of geomagnetic fields has been surrounded by a fog of mystery, but recent developments, mostly due to the work of Sir Edward Bullard in England and W. M. Elsasser in the United States, strongly suggest that these fields are caused by the nonuniform rotation of different layers in the earth's fluid and no doubt highly ionized core, which occupies about 60 percent of the radius of our globe. In fact, it is natural to assume that the heat conduction in the core of the earth is mostly due to convective currents. Because of the earth's rotation around its axis, the rising convective currents will be deflected *westward*, and we would expect that the outer layers of the molten iron core would be rotating more slowly than its inner layers.

FIG. 32-10 A self-exciting disk dynamo.



How can such a differential rotation of the electrically conductive iron core give rise to a magnetic field? A possible way is represented schematically in Fig. 32-10, which is a "self-exciting disk dynamo." It consists of a rotating metal disk *A* and a ring-shaped conductor *B* (which may have many windings), connected by brushes *C* and *C′* with the disk and the axle, respectively. Imagine that a weak magnetic field is applied, let us say upward, parallel to the axle. According to the law applicable to a conductor moving through a magnetic field, an electric current will be induced in the disk and will flow radially from the periphery to the axle. Flowing throught the ring-shaped conductor, this current will increase the strength of the magnetic field, which will, in turn, increase the strength of the current, etc. Thus a very weak magnetic field applied to this system will rapidly grow to an upper limit determined by the electric resistivity of the metallic parts. These generators are impractical for technical applications, but they may represent a model of what happens inside the earth, where the inner part of the conductive iron core rotates with respect to its outer part. Calculations indicate that such a process might lead to magnetic fields strong enough to account for the main magnetic field surrounding the earth.

32-8 Physics of the Atmosphere

As we rise above the earth's surface, the temperature of the air decreases with increasing altitude at a rate of about 6°C/km. This decrease is due to the fact that *atmospheric air is not heated directly by the sun's rays, but gets its heat through the convective currents that rise from the warmed surface of the earth.* As the air that is heated by a direct contact with the ground rises, it expands because the pressure at high altitudes is less than that at low altitudes. We know that such an expansion inevitably causes a decrease in temperature. In fact, the observed temperature gradient of 6°C/km is in good agreement with the cooling that we would expect to result from the expansion of the ascending air masses.

Up to the beginning of the present century, it was believed that the decrease of air temperature with altitude continued all the way up to the upper fringes of the atmosphere, the temperature of which was expected to approach the nearly absolute zero temperature of interplanetary space. Subsequent studies, carried out by means of balloons and more recently by rockets, indicate, however, that this is not true. It has been found that the decrease of air temperature continues only up to an altitude of about 20 km, where it reaches a minimum of about 210°K (−60°C or −76°F). At still higher altitudes, the temperature begins to rise again, reaching a value close to the freezing point of water, but it then drops down to about 180°K (−90°C or −130°F) at an altitude of 80 km (Fig. 32-11). The vertical convective currents stop short of the altitude of 20 km, and from 20 to 80 km there is almost no

mixing at all between the layers (strata) of air lying above each other. The lower, convective layer of the atmosphere is called the *troposphere.* The moistureless and cloudless air layer above 20 km is known as the *stratosphere.*

Above 80 km, the temperature changes are reversed again, and the temperature begins to increase with increasing altitudes, reaching room temperature at about 130 km, the temperature of boiling water at 160 km, and the temperature of molten lead at 250 km! (See Fig. 32-11.) You should not think, however, that this means that if you ascended to this altitude you would be roasted alive. Although the air molecules there have the velocity they would have at these high temperatures on the earth's surface, the density of the air becomes extremely low (less than one billionth of its standard density) and its ability to conduct heat to or from material bodies becomes negligibly small. In a warm or a cold room on the ground, each square centimeter of the surface of your body receives some 10^{25} molecular impacts every second, and this enormous number of impacts heats or cools our body very rapidly. At an altitude of 250 km, however, the number of molecular impacts is reduced by a factor of many billions, and the amount of heat the molecules can communicate or take away from the body becomes correspondingly lower. The temperature of any material object, such as an ascending rocket or an artificial satellite, at these altitudes is completely determined by the absorption of the sun's rays and by the reemission of the absorbed energy in the form of heat radiation.

The increase of the temperature of the air in the upper atmosphere is due to the strong absorption of the ultraviolet radiation of the sun. When the energy-loaded ultraviolet-light quanta from the sun enter the upper fringes of the atmosphere, they kick off electrons from the outer shells of the nitrogen and oxygen atoms, give them high velocities, and thus maintain a high degree of thermal agitation. Because of the extremely low densities and the correspondingly low collision probabilities at these alitituds, the positive ions and the electrons travel for a long time before they have a chance to recombine into neutral atoms. As a result, *the air in this region is continuously maintained at a high degree of ionization.* This region of the terrestrial atmosphere, the ionosphere, possesses a high degree of electrical conductivity (because of the free electrons and positive ions) and is a good reflector for radio waves, which can propagate all the way around the globe by bouncing between the earth's surface and the reflecting layers of the ionosphere (Fig. 32-12). Another consequence of the high electrical conductivity of the ionosphere is the presence in it of electric currents that flow through it in a somewhat irregular fashion. The variable magnetic fields produced by these ionospheric currents add to the permanent magnetic field of the earth, to cause minor fluctuations in it.

In addition to the ultraviolet radiation from the sun, which is almost completely absorbed in the ionosphere and gives the latter its very high

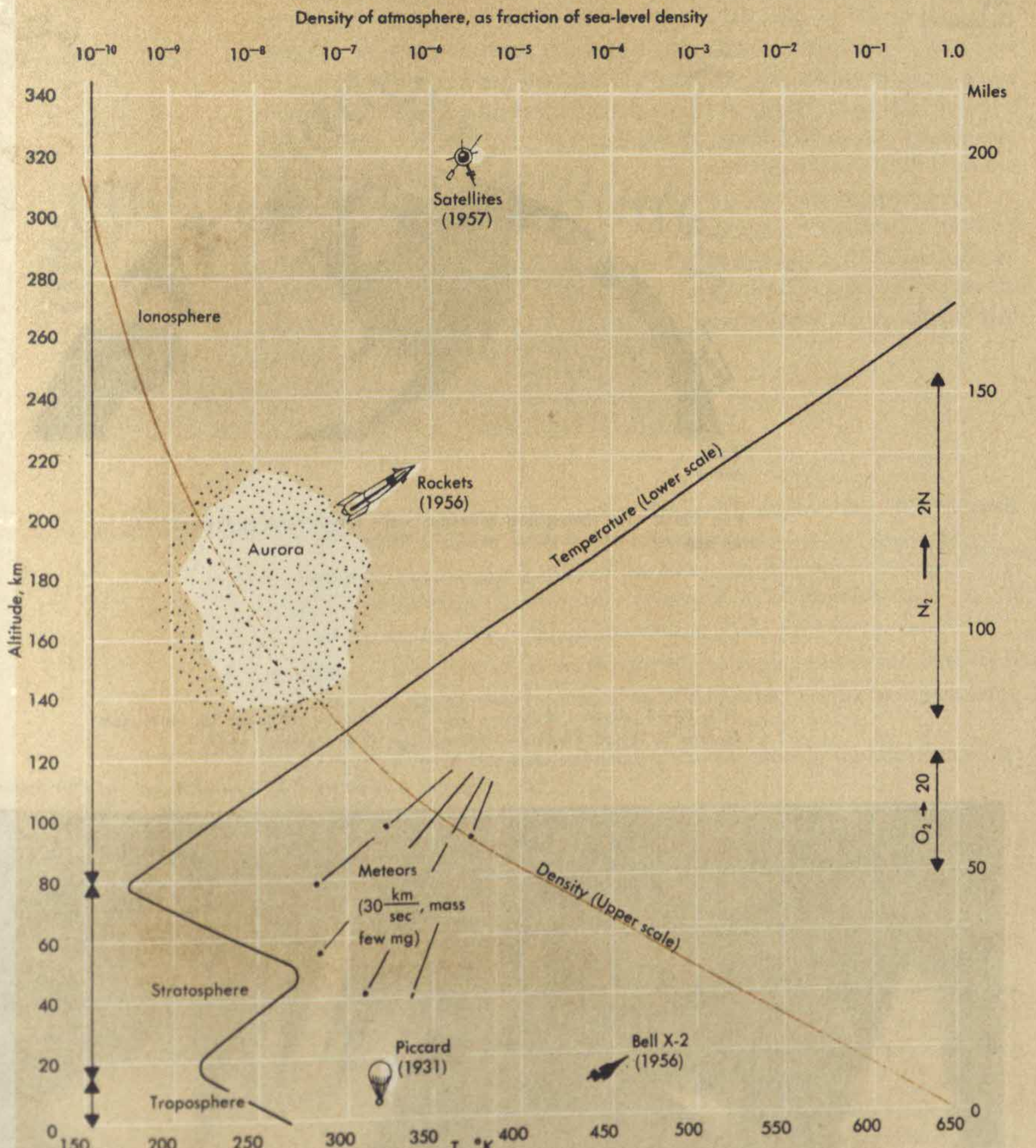


FIG. 32-11 Density and temperature of the earth's atmosphere at high altitudes.

FIG. 32-12 The ionosphere reflecting radio waves back to the earth's surface.

FIG. 32-13 Auroral draperies over Oslo, Norway, October 13, 1916. From George F. Taylor, *Elementary Meteorology*, (Prentice-Hall, 1954).

Courtesy U.S. Weather Bureau and Carl Stormer.

temperature, this space frontier of our atmosphere is subject to continuous bombardment by other high-energy radiations. It is penetrated by streams of comparatively slow-moving electrically charged particles that come from the sun and are deflected by the magnetic field of the earth to the poles, where they produce the magnificent phenomenon of the aurora (Fig. 32-13). It is pierced by multibillion-volt cosmic-ray particles that produce all kinds of secondary radiations in it. It is visited by meteors that shoot in at high speed from interplanetary space.

It is a matter of some concern that astronauts, occasionally orbiting around the earth in their artificial satellites, be properly shielded from the effects of these radiations and particles. Later, when men live for months on the airless moon or in an orbiting observatory, the problem will become even more important.

Questions

(32-1)

1. Assuming the density of the rock to average 3.0 gm/cm^3, show that the pressure at a depth of 50 km deep in the earth is about 1.5×10^4 atmospheres.

2. Assuming the average density of the rock to be 3.1 gm/cm^2, at about what depth below the earth's surface is the pressure 2×10^4 atmospheres?

3. About what fraction of the earth's volume is occupied by the core?

4. About what fraction of the earth-core volume is occupied by the inner core?

5. Within the earth's mantle, does the speed with which seismic waves are transmitted increase or decrease with depth? (Refer to Fig. 32-3.)

6. Is the speed of transmission of seismic waves greater in the core than in the mantle, or less? (See Fig. 32-3.)

(32-5)

7. A large floating ice field protrudes 10 ft, on the average, above the surface of the sea water on which it floats. (Density of sea water = 1.03 gm/cm^3; of ice, 0.92 gm/cm^3.) How far below the water surface does the ice field extend?

8. On the ice field of Question 7, a considerable area stands 15 ft above the surrounding ice. Beneath this area, how far does the ice extend downward into the water beyond the general level of the bottom of the ice field?

9. Greenland is covered with a sheet of ice (density = 0.9 gm/cm^3) with an average thickness of 2500 ft. Consider Greenland as a rigid island floating on plastic material of density 3.0 gm/cm^3. How far would Greenland rise if all its ice were to melt?

10. Consider a hypothetical continent of granitic rock of density 2.70 gm/cm^3. Its surface stands at an average elevation of 0.5 mile above the surrounding sea, which is 3 miles deep. Beneath the sea and the continent is plastic rock whose density is 3.2 gm/cm^3. How thick is the continent?

(32-6)

11. Confirm (approximately) the shrinkage of the earth's circumference by cooling, as given in Sec. 32-6.

12. Show that a cubic mile of material which changed in density from 5.5 to 9.5 gm/cm^3 would thereby shrink 0.42 mi^3 in volume, to become 0.58 mi^3.

chapter / thirty-three

Astrophysics

33-1 Planetary Atmospheres

We on earth are surrounded by an extraordinary atmosphere that has to a large extent been created by the organisms that live in it. Plants use the carbon dioxide in the air and exhale oxygen as a useless waste product. Animals use the oxygen and exhale carbon dioxide as waste. It is generally conceded that, without the ceaseless work of plant life, the earth's atmosphere would contain only negligible amounts of oxygen, and animal life would be impossible. Speculations of this sort lie within the provinces of the biologist, the biochemist, and the geochemist. The physicist must explain why the earth has as much atmosphere as it does, while Mercury and the moon have none, and planets like Saturn and Jupiter have heavy atmospheres that are thousands of miles deep.

Earlier in the book we talked of escape velocity, and derived for it the expression

$$v = \sqrt{\frac{2GM}{r}}$$

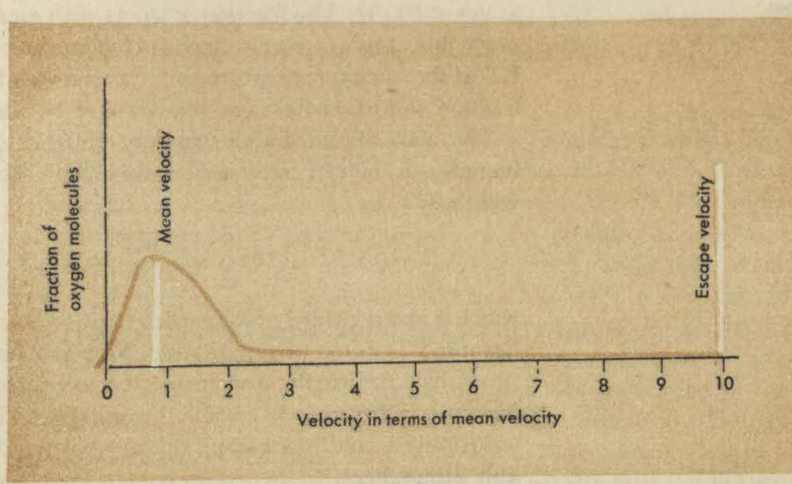


FIG. 33-1 Maxwell's distribution of velocities of oxygen atoms at 600°K.

in which G is the universal gravitational constant, 6.67×10^{-8} CGS units, M is the mass of the planet in grams, and r its radius in centimeters. The expression was worked out for projectiles or spaceships; but the fact that m, the mass of the projectile, does not appear shows that the formula will apply equally well to a gas molecule. If an atom of oxygen or nitrogen near the top of the earth's atmosphere happens to be heading out toward space with a speed greater than earth's escape velocity of 1.12×10^6 cm/sec, or 6.96 mi/sec, the earth will lose that molecule.

Let us figure how the speed of high-altitude oxygen compares with this escape velocity. In the ionosphere, the oxygen molecule has been broken down into oxygen atoms at an estimated average temperature, let us say, of 600°K. We know that at ordinary temperatures (300°K) at the earth's surface, an oxygen molecule travels approximately 500 m/sec, or 5×10^4 cm/sec. An oxygen *atom* has half the mass of the molecule, and at the same temperature would have the same kinetic energy, which means its speed would be $\sqrt{2} \times 5 \times 10^4 = 7 \times 10^4$ cm/sec. In the ionosphere, at 600°K, we have doubled the temperature and hence have doubled the kinetic energy. To double the KE, the speed must again be multiplied by $\sqrt{2}$, which tells us that an oxygen atom at 600°K will have an average speed of $\sqrt{2} \times 7 \times 10^4 = 10^5$ cm/sec. This average speed is about 0.1 of the escape velocity, and the fraction of oxygen atoms having escape velocity will be very small indeed. Figure 33-1 shows a Maxwellian distribution of velocities for oxygen

atoms at 600°K. The fraction of atoms with escape velocity is completely negligible. Lighter atoms such as hydrogen or helium, with the same KE at the same temperature, will travel much faster, and a much larger fraction will exceed escape velocity and be able to leave the earth.

The mass of the planet Mercury is 0.045 of the earth's mass, and its radius is 0.39 of the earth's radius. The escape velocity of Mercury will thus be

$$1.12 \times 10^6 \times \sqrt{\frac{0.045}{0.39}} = 3.8 \times 10^5 \text{ cm/sec}$$

which is about a third of the earth's escape velocity of 1.12×10^6 cm/sec. Because it is much nearer the sun, Mercury has much higher temperatures than the earth, and hence the average speeds of any atoms of Mercurian atmosphere would be greater. The combination of high temperatures and low escape velocity has left Mercury with no observable atmosphere.

The large, massive planets, Jupiter and Saturn, have high escape velocities and low temperatures; hence they have kept enormous atmospheres, including even the light gas hydrogen, which the earth has been unable to hold.

33-2 Stellar Atmospheres

We have spoken before of the spectrum of sunlight, and how it is produced in the solar atmosphere. The sun, since it is made entirely of gas, has no definite boundary as the solid earth has. At a certain depth in the solar atmosphere, however, the atoms are compressed so closely together that they emit a continuous spectrum of all frequencies, rather than the definite spectral lines that come from an excited gas at lower pressure. This layer, the source of most of the sun's visible light, is called the *photosphere* and is generally taken as the sun's boundary.

The temperature of the sun's photosphere can be estimated by a study of the energy distribution in the continuous spectrum of the sun. In fact, we have seen that the wavelength corresponding to the maximum intensity in the spectrum decreases in inverse proportion to the absolute temperature (Wien's law). In the observed solar spectrum, the maximum intensity is at the wavelength 4700 Å, from which we deduce that the surface temperature of the sun is about 6000°K. An alternative method is based on the Stefan-Boltzmann relation between the temperature of the surface and the amount of radiation it emits, and it leads to about the same value.

When the same methods are applied to the analysis of the spectra of more distant stars, we find that these bodies have surface temperatures

ranging from a dull red of a thousand degrees or so to temperatures in the tens, or even hundreds, of thousands of degrees.

33-3 Other Spectroscopic Information

Astronomers have been interested for a long time in knowing whether the sun possesses a magnetic field similar to that of the earth. Since, of course, no one can get to the sun with a magnetic needle, the presence of that magnetic field could be revealed only through its influence on light emitted by the sun. Laboratory experiments show that the line spectra of elements are influenced by magnetic fields. If, for example, we place a bunsen burner flame containing some table salt between the poles of a strong electromagnet, the yellow lines of the sodium spectrum will split into several components, and the separation of these components will increase in proportion to the magnetic field. This so-called *Zeeman effect* is caused by the action of the magnetic field on the orbital motion of atomic electrons and can be used for a direct measurement of the intensity of that field. The application of this method to the study of magnetic fields on the surface of the sun was initiated in 1908 by an American astronomer, G. E. Hale, and it led to many important discoveries. While it seems that the sun does not possess a regular magnetic field similar to that of the earth, very strong magnetic activity is associated with sunspots, flares, prominences, and other irregular features characteristic of the solar surface. We know that the regular magnetic field of the earth is occasionally disturbed by what we call "magnetic storms"; there seem to be nothing but violent magnetic storms on the surface of the sun!

It is very difficult to measure the motion of stars in a direction at right angles to our line of sight. The distances are so enormous that only the closest stars show any apparent motion against the background of more distant stars, even when photographs taken many years apart are compared. The motion of stars along the line of sight—that is, approach or recession—is quite a different matter, however. The Doppler effect, which we have already mentioned in connection with sound, also works with light waves and tells how fast the source of light is approaching or receding from us. The equation we used for the Doppler effect in sound was

$$f' = f \times \frac{v \pm v_o}{v \pm v_s}.$$

This equation applies to a wave traveling with a speed v through a definite medium—air, usually, or water, or some other material substance. It could apply in this form to the prerelativistic theories of a

material ether, but these were discarded more than 60 years ago. An accurate relativistic expression for the Doppler effect is mathematically a little too much for us to tackle, but if we keep v_o and v_s small compared with the speed of light c, things can be considerably simplified, and the following expression can be derived:

$$v_{\mathrm{rel}} = \frac{c\Delta\lambda}{\lambda}.$$

We can obtain, of course, only the *relative* speed between source and observer, v_{rel}, since there is no such thing as an absolute motion of either source or observer. The shift in apparent wavelength of a spectral line compared with the same line photographed from a stationary laboratory source is $\Delta\lambda$.

For example, a stellar spectrum may be photographed on the same plate and with the same spectroscopic equipment as the spectrum of certain comparison elements in the laboratory. One hydrogen line has a wavelength of 6562.8 Å, and if the corresponding Fraunhofer line in the spectrum of a star falls at the location corresponding to 6564.2 Å, we have the wavelength shifted by 6564.2 − 6562.8 = 1.4 Å. This shift gives us for the relative velocity of the star,

$$v = \frac{3 \times 10^5 \times 1.4}{6563} = 64 \text{ km/sec}.$$

Since the spectral line from the star has been shifted to a longer wavelength, which means a lower frequency, we know that this computed speed is a recession away from the earth.

33-4 Properties of Matter Inside the Sun

Direct observations of the sun are limited to the surface of the photosphere, and no human eye can penetrate into its deep interior. Yet we know more about the central regions of the sun than about the core of the earth, in spite of the fact that the earth is right under our feet and the sun is a hundred million miles away! The reason for this strange fact is that in both cases we have to make conclusions on the basis of theoretical considerations which require knowledge of the properties of matter under the physical conditions existing in the deep interior of these two celestial bodies. In the case of the earth, we have to deal with the physical properties of a molten substance subjected to extremely high pressures inaccessible in our laboratories. Being unable to study these properties experimentally, we also lack any reliable theoretical predictions concerning them, since, as was mentioned earlier in the book, the theory of the liquid state of matter is extremely difficult. It would seem that in the case of the sun's interior, where the temperature and pressure exceed, by a large margin, those encountered in the interior of the earth, the

situation would be even worse. But this is not so. It turns out, in fact, that just because of these tremendously high temperatures which exist the solar interior, the properties of matter become very simple and easily predictable on the basis of our present knowledge concerning atomic structure. The reason for this is that, *under the conditions existing in the solar interior, all molecules and atoms which form ordinary matter are almost completely broken up into their constituent parts.*

Consider a glass of water that is brought gradually to a higher and higher temperature. At room temperature, we have a complicated structure in which water molecules composed of hydrogen and oxygen atoms are held together by intermolecular cohesive forces. With the increase of the temperature, the molecules will be torn apart and we obtain water vapor, in which individual molecules fly freely through space and very rarely collide with one another. The physical properties of water vapor, being subject to classical gas laws, are much simpler to describe than those of liquid water. However, the properties of individual molecules (such as an absorption spectrum) are still almost as complicated as those in the liquid state. When the temperature rises to still higher values, violent thermal collisions between the water vapor molecules will break them up (*thermal dissociation*) into individual atoms of hydrogen and oxygen. In this state, the optical properties of the gas become more predictable, since they pertain now to individual atoms of hydrogen and oxygen, which are simpler than the composite molecules of water. Now, if the temperature goes still higher, the increasing violence of thermal collisions will begin to strip the individual atoms of their electronic shells (*thermal ionization*). At the temperature of 6000°K that prevails on the surface of the sun, most hydrogen atoms are broken up into protons and electrons, while the atoms of oxygen are stripped of two or three electrons from their outer shells. At the still higher temperatures that are encountered in the solar interior, oxygen atoms, as well as the atoms of all other heavier elements, will be almost completely stripped of all their electron shells, and there will be a *mixture of bare nuclei and free electrons involved in a violent thermal motion.*

This situation simplifies the picture quite a bit. First of all, we can expect that no matter how high the density, solar matter can be considered as an *ideal gas*. Indeed, as we have seen, an ideal gas is characterized by the smallness of the individual particles as compared with their relative distances. When a material consists of atoms and molecules which are about 10^{-8} cm in diameter, its particles must be packed tightly together when its density is about that of water. But once the atoms are broken into nuclei and free electrons, which have respective diameters of only 10^{-12} to 10^{-13} cm, the situation becomes entirely different, and the properties of gas will be retained up to much higher

densities. In fact, in order to make completely ionized matter deviate from the laws of ideal gas, we have to compress it to densities at which the electrons and the nuclei are squeezed together as are sardines in a can. This density exceeds that of water by a factor of $(10^{-8}/10^{-12})^3 = 10^{12}$ gm/cm^3, and is far beyond the densities that are encountered inside any ordinary star.* Thus, even though the material in the center of the sun is at 100 times the density of water, it can still be considered as a perfect gas, obeying the simple gas laws formulated previously.

Knowing the physical conditions (temperature, density, etc.) on the surface of the sun and the laws governing the gaseous material in its interior, we can calculate, step by step, the change of these conditions as we proceed below the solar surface and toward its center. Such calculations were first carried out by the famous British astronomer Arthur Eddington, and gave us the first insight into the inside of our sun as well as of the other stars. The central temperature of the sun proves to be in the neighborhood of 20,000,000°K!

33-5 Energy Production in the Sun

A primitive man, as well as a modern man with little education in the field of science, would consider the sun as "burning" in very much the same way as logs burn in a fireplace or as kerosene burns in a lamp. However, it is easy to show that, even if it were made of the best rocket-fuel mixture, the burning sun would not last for more than a few thousand years.

In the middle of the last century, shortly after the discovery of the law of equivalence of mechanical energy and heat, a new theory of solar energy production was proposed simultaneously by H. von Helmholtz in Germany (1821–1894) and Lord Kelvin in England (1824–1907). They indicated that, being a giant sphere of hot gas which loses its energy through radiation from the suface, the sun is bound to contract, though very slowly. During this contraction, the forces of gravity acting between different parts of the solar body do mechanical work that is then turned into heat, according to the equivalence principle of mechanical and thermal energy. It has been calculated that, being much more powerful than chemical reactions (such as burning), this process could maintain the sun at its present luminosity for a couple of hundred million years. At the time of Kelvin and Helmholtz, a few hundred million years seemed to be a sufficiently long period to explain the events of historical geology, and so their point of view was accepted without objection. We know now, however, that life on the earth must have existed for at least half a billion years, and the solid crust and the oceans must

* Recently, astrophysicists have begun to seriously speculate about the possible existence of "neutron stars," in which the pressure and density are so great that protons and electrons are forced to combine into neutrons.

be much older still. Thus the contraction of the sun certainly could not have provided enough energy to maintain the surface of the earth and its oceans comfortably warm for the origin and evolution of living organisms.

The discovery of radioactivity and the recognition of the fact that the energy stored within atomic nuclei exceeds, by a factor of a million, the energy liberated in ordinary chemical transformations, threw an entirely new light on the problem of solar-energy sources. If the sun could have existed for several thousands of years fed by an ordinary chemical reaction (burning), nuclear energy sources are surely rich enough to supply an equal amount of energy for billions of years. The trouble is, however, that in natural radioactive decay, the liberation of nuclear energy is extremely slow. In order to explain the observed mean rate of energy production in the sun (2 erg/gm-sec), we would have to assume that the sun is composed almost entirely of uranium, thorium, and their decay products. Thus we are forced to the conclusion that the liberation of nuclear energy inside the sun is not an ordinary radioactive decay, but rather some kind of induced nuclear transformation caused by the specific physical conditions in the solar interior. It is natural to expect that the factor responsible for the induced nuclear transformations is the tremendously high temperature existing in the solar interior. Indeed, at a temperature of 2×10^{7}°K, the kinetic energy of thermal motion is 4.2×10^{-9} erg per particle, which, expressed in electric units, amounts to 3 Kev. This is considerably smaller than the energies ordinarily used in the experiments on nuclear bombardment (1 Mev and up), but we must take into account that, whereas artificially accelerated nuclear projectiles rapidly lose their initial energy and have only a small chance to hit the target nucleus before coming out of the game, thermal motion continues indefinitely and the particles involved in it collide with each other for hours, centuries, and billions of years. The calculations carried out in this direction in 1929 by the Austrian physicist F. G. Houtermans and the British astronomer R. Atkinson led to the conclusion that at the temperatures existing in the solar interior, thermonuclear reactions between hydrogen nuclei (protons) and the nuclei of other light elements can be expected to liberate sufficient amounts of nuclear energy to explain the observed radiation of the sun.

33-6 Carbon Cycle and H-H Reaction

Although the work of Houtermans and Atkinson proved, beyond any doubt, that the energy production inside the sun is due to thermonuclear reactions between hydrogen and some light elements, the exact nature of these reactions remained obscure for another decade because of the lack of experimental knowledge concerning the result of nuclear bombardment by fast protons. However, with the pioneering work of

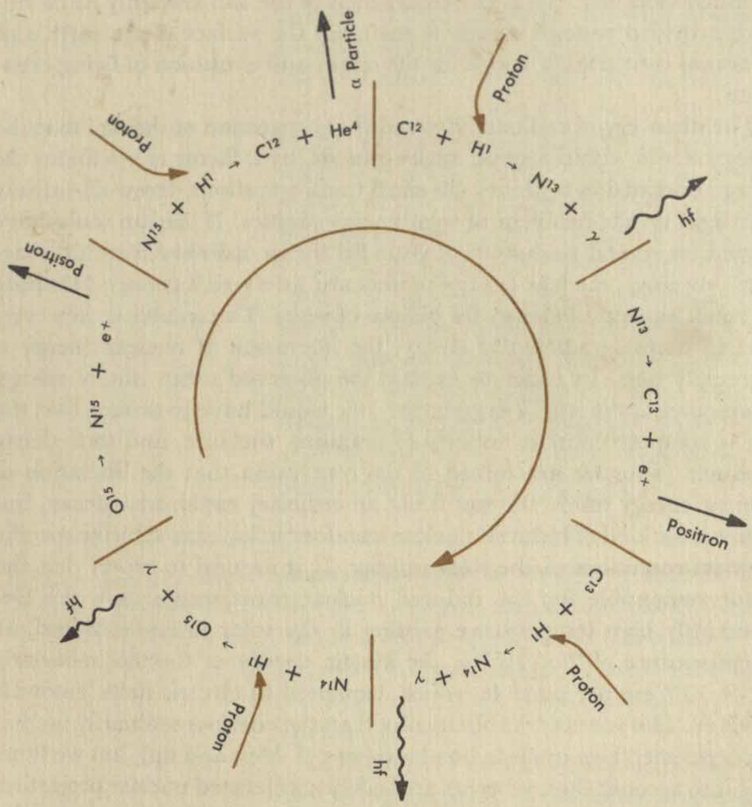


FIG. 33–2 Diagram of the nuclear reactions in the carbon cycle, responsible for most of the energy production of massive stars.

Cockcroft and Walton on artificially accelerated proton beams and subsequent work in this direction, enough material was collected in this field to permit the solution of the solar-energy problem. One possible solution was proposed in 1937 by H. Bethe in the United States and C. von Weizsacker in Germany (independently of one another) and is known as the *carbon cycle*, while the other possibility was conceived by an American physicist, Charles Critchfield, and is known as the *H-H reaction*. The net total result of both reactions is the transformation of hydrogen into helium, but it is achieved in a different manner in each reaction.

In carbon cycles, the atom of carbon can be considered as a "nuclear catalyst" that helps unite 4 independent protons into a single α particle by capturing them one by one and holding them together until the union is achieved. After 4 protons are caught and the newly formed α particle

is released, we get back the original carbon atom which can go again through the next cycle. The series of reactions constituting the carbon cycle is shown in Fig. 33-2. At the temperature and pressure at the center of the sun, the total period of the cycle is 6×10^6 years, and the total energy liberated by the cycle is 4×10^{-5} erg, which yields for the rate of energy liberation per carbon atom the value of 2×10^{-19} erg/sec. Since, according to present data concerning the chemical composition of the sun, each gram of solar material contains 10^{-4} gm of carbon (5×10^{18} carbon atoms), this leads to the total rate of energy liberation of 1 erg/gm/sec, which is only 1 percent of the observed rate of energy production in the small hot core of the sun, where these reactions can take place.

In the H-H reaction, the first step is the formation of a deuterium nucleus in the process of thermal collision between two protons:

$${}_1H^1 + {}_1H^1 \longrightarrow {}_1D^2 + e^+ + \text{neutrino}.$$

This is followed by a series of thermonuclear reactions which build up the so-formed deuteron into an α particle:

$${}_1D^2 + {}_1H^1 \longrightarrow {}_2He^3 + \gamma \text{ ray}$$

$${}_2He^3 + {}_2He^3 \longrightarrow {}_2He^4 + 2{}_1H^1.$$

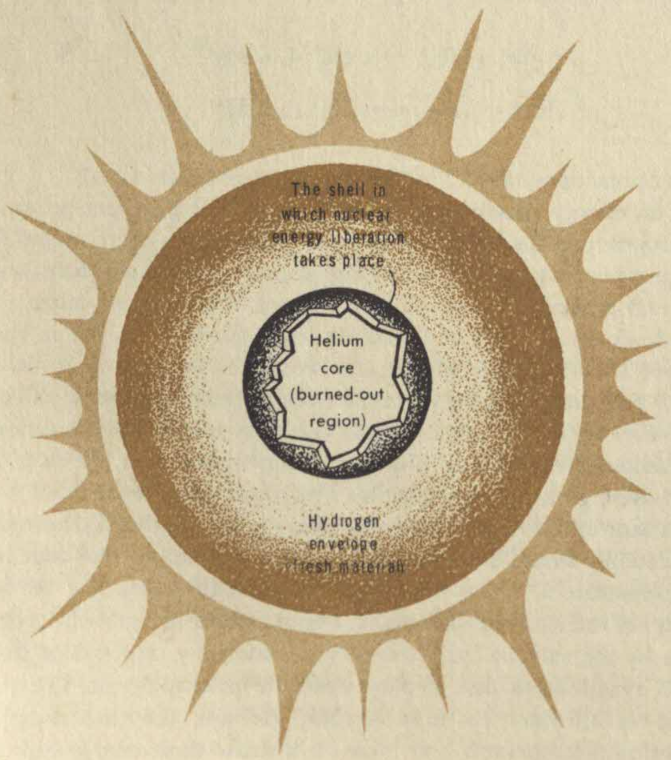
This series of reactions takes 3×10^9 years and liberates 4×10^{-5} erg, leading to an energy production rate of 5×10^{-22} erg/sec per proton. Since hydrogen constitutes about 50 per cent of the solar material (2×10^{23} protons per gram), the total rate of energy liberation comes out to be 100 ergs/gm/sec, in good agreement with the observed value.

The predominance of the H-H reaction over the carbon cycle in the sun, however, is not the general rule, and is reversed in many other stars. The point is that these two sets of thermonuclear reactions possess different sensitivities to temperature. While the rate of the H-H reaction increases comparatively slowly with increasing temperature, the rate of the carbon cycle goes up very rapidly. The more massive the star, the hotter it is inside and the more violent are the energy-producing thermonuclear reactions. Thus, if we take Sirius, for example, we find that its central temperature is 2.3×10^7°K (compared with only 2.0×10^7°K in the center of the sun). At this higher temperature, the rate of energy production by the carbon cycle becomes considerably larger than that of the H-H reaction, so that it plays here the principal role. On the other hand, in all stars less massive than the sun, the carbon cycle becomes quite unimportant, and these stars draw their energy almost entirely from the H-H reaction.

33-7 The Future of Our Sun

Since the energy radiated by the sun is due to the continuous transformation of hydrogen into helium in its interior, the sun evidently cannot shine for an eternity, and is bound to run out of fuel sometime in the future. It is estimated that during the 5 billion years of its existence, the core of our sun has used about half of its original supply of hydrogen, so that it still has enough nuclear fuel for another 5 billion years. What will happen 5 billion years from now, when our sun comes close to the end of its resources? To answer this question, we have to remember that thermonuclear reactions proceed almost exclusively near the center of the sun, where the temperature is the highest. Thus the shortage of nuclear fuel will be felt first in the central regions of the sun, where all the originally available hydrogen will have been transformed into helium. We can easily visualize that this will result in a rearrangement of things in the solar interior in such a way that the high-temperature region will move to the interface between the "burned-out core" and the outer layers that still contain enough hydrogen to maintain a nuclear

FIG. 33–3 The shell-source model for giant red stars.



fire. The internal structure of the sun, therefore, will be transformed from a so-called *point-source model* (energy source in the center) to a *shell-source model*, in which the energy is liberated in a thin spherical shell that separates the burned-out core from the rest of the solar body (Fig. 33-3). As more and more hydrogen is consumed, the "shell" will move outward from the center, as does a ring of fire that has been started by a carelessly dropped match in a dry grass field.

It was suggested by Charles Critchfield and George Gamow, and later confirmed by the more detailed calculations of M. Schwarzschild and his associates, that the formation of such a shell source inside the sun (or any other star) must result in a steady growth of the star's size and in a gradual increase of its luminosity. In fact, within a few hundred million years after the shell source is formed, the diameter of the sun is expected to become as large as the orbit of Venus, and its luminosity will increase by a factor of between 10 and 20, making the oceans on the earth boil violently. After this last effort, the sun will begin to shrink and fade out again, until it becomes quite faint and insignificant. But there is no reason for immediate panic—we still have 5 billion years to go!

Questions

(33-1)

1. (a) What will be the average (rms) speed of a hydrogen atom at a temperature of 600°K? (b) What fraction is this of its escape velocity?
2. At a temperature of 327°C, what is the approximate rms speed of an atom of helium? How does this compare with its escape velocity?
3. The temperature of the sun's chromosphere (Fig. 24-6) runs as high as 50,000°K. The sun's radius is 100 times as large as the earth's, and its mass is greater by a factor of about 3×10^5. In the chromosphere, how does the rms speed of a hydrogen atom compare with its escape velocity?
4. Jupiter has 318 times the mass of the earth, and its radius is 11 times as great. (a) What is the escape velocity from Jupiter? (b) What would the temperature of Jupiter have to be in order for oxygen atoms to have one-third of this escape velocity? (Actually, the surface temperature of Jupiter is more than 200°C below zero.)

(33-2)

5. Use Wien's law to find the most intense wavelength in the radiation from a star with a surface temperature of 40,000°K.
6. What is the wavelength of the most intense radiation from a star whose surface temperature is 3000°K?

(33-3)

7. A line in the spectrum of a star has a wavelength of 5912.6 Å; it is identified as a line whose laboratory wavelength is 5920.0 Å. (a) Is the relative velocity of the star toward or away from the earth? (b) What is the star's radial velocity?

8. The wavelengths of all the lines in the spectrum of a distant galaxy (actually, the combined spectra of the billions of stars comprising the galaxy) are 0.5 percent longer than the corresponding laboratory lines. (a) Is this galaxy approaching or receding? (b) What is its radial velocity relative to the earth?

9. The equatorial regions of the sun make approximately one revolution in 25 days, and the solar diameter is about 1.4×10^6 km. By how many Ångstroms will the apparent wavelength of a 6500-Å line be shifted when the observed light comes from the limb of the sun moving toward the earth?

10. The spectrum of a star is photographed twice: once when the star is near the horizon after sunset, and again when it is just above the horizon before dawn. By how much will the wavelengths differ in these two spectra, for a line whose laboratory wavelength is 4987 Å?

(33-6) 11. About 4 million metric tons of energy per second are radiated from the surface of the sun. How many protons per second must be converted into helium to produce energy at this rate?

Conversion Tables

LENGTH

CGS		MKS		Other
1 cm	=	10^{-2} m	=	0.3937 in.
10^2 cm	=	1 m	=	39.37 in.
10^5 cm	=	10^3 m	=	1 km
10^{-4} cm	=	10^{-6} m	=	1 μ
2.54 cm	=	0.0254 m	=	1 in.
10^{-8} cm	=	10^{-10} m	=	1 Å
1.609×10^5 cm	=	1609 m	=	1 mile

AREA

CGS		MKS		Other
1 cm^2	=	10^{-4} m^2	=	0.1550 in.^{2}
10^4 cm^2	=	1 m^2	=	10.76 ft^2
6.452 cm^2	=	6.452×10^{-4} m^2	=	1 in.^{2}
929 cm^2	=	0.0929 m^2	=	1 ft^2

VOLUME

CGS		MKS		Other
1 cm^3	=	10^{-6} m^3	=	0.06102 $in.^3$
10^6 cm^3	=	1 m^3	=	35.31 ft^3
16.39 cm^3	=	1.639×10^{-5} m^3	=	1 $in.^3$
2.832×10^4 cm^3	=	0.02832 m^3	=	1 ft^3

MASS

CGS		MKS		Other
1 gm	=	10^{-3} kg	=	2.205×10^{-3} lb
10^3 gm	=	1 kg	=	2.205 lb
453.6 gm	=	0.4536 kg	=	1 lb
1.459×10^4 gm	=	14.59 kg	=	1 slug
1.6597×10^{-24} gm	=	1.6597×10^{-27} kg	=	1 amu = 931 Mev

ENERGY

CGS		MKS		Other
1 erg	=	10^{-7} joule	=	7.376×10^{-3} ft-lb
10^7 erg	=	1 joule	=	0.7376 ft-lb
4.186×10^7 erg	=	4.186 joule	=	1 calorie
3.6×10^{13} erg	=	3.6×10^6 joule	=	1 KWH
1.602×10^{-6} erg	=	1.602×10^{-13} joule	=	1 Mev = 1.074×10^{-3} amu

ELECTRICAL AND MAGNETIC

CGS		MKS
1 esu	=	3.336×10^{-10} coul
2.998×10^9 esu	=	1 coul
10^8 maxwell	=	1 weber
10^4 gauss	=	1 weber/m^2
1 gauss	=	10^{-4} weber/m^2

MISCELLANEOUS RELATIONSHIPS

60 mi/hr = 88 ft/sec
1 radian = 57.30 degrees
1 horsepower = 746 watts
1 lb/in.^{2} = 6.895×10^4 dyne/cm^2 = 6.895×10^3 newtons/m^2
1 atm = 1.013×10^6 dyne/cm^2 = 76 cm Hg = 14.70 lb/in.^{2}

SINES, COSINES, AND TANGENTS

Angle	Sine	Cosine	Tangent
0°	0.000	1.000	0.000
1°	.017	1.000	.017
2°	.035	0.999	.035
3°	.052	.999	.052
4°	.070	.998	.070
5°	.087	.996	.087
6°	.105	.995	.105
7°	.122	.993	.123
8°	.139	.990	.141
9°	.156	.988	.158
10°	.174	.985	.176
11°	.191	.982	.194
12°	.208	.978	.213
13°	.225	.974	.231
14°	.242	.970	.249
15°	.259	.966	.268
16°	.276	.961	.287
17°	.292	.956	.306
18°	.309	.951	.325
19°	.326	.946	.344
20°	.342	.940	.364
21°	.358	.934	.384
22°	.375	.927	.404
23°	.391	.921	.424
24°	.407	.914	.445
25°	.423	.906	.466
26°	.438	.899	.488
27°	.454	.891	.510
28°	.469	.883	.532
29°	.485	.875	.554
30°	.500	.866	.577
31°	.515	.857	.601
32°	.530	.848	.625
33°	.545	.839	.649
34°	.559	.829	.675
35°	.574	.819	.700
36°	.588	.809	.727
37°	.602	.799	.754
38°	.616	.788	.781
39°	.629	.777	.810
40°	.643	.766	.839
41°	.656	.755	.869
42°	.669	.743	.900
43°	.682	.731	.933
44°	.695	.719	.966
45°	.707	.707	1.000
46°	0.719	0.695	1.036
47°	.731	.682	1.072
48°	.743	.669	1.111
49°	.755	.656	1.150
50°	.766	.643	1.192
51°	.777	.629	1.235
52°	.788	.616	1.280
53°	.799	.602	1.327
54°	.809	.588	1.376
55°	.819	.574	1.428
56°	.829	.559	1.483
57°	.839	.545	1.540
58°	.848	.530	1.600
59°	.857	.515	1.664
60°	.866	.500	1.732
61°	.875	.485	1.804
62°	.883	.469	1.881
63°	.891	.454	1.963
64°	.899	.438	2.050
65°	.906	.423	2.145
66°	.914	.407	2.246
67°	.921	.391	2.356
68°	.927	.375	2.475
69°	.934	.358	2.605
70°	.940	.342	2.747
71°	.946	.326	2.904
72°	.951	.309	3.078
73°	.956	.292	3.271
74°	.961	.276	3.487
75°	.966	.259	3.732
76°	.970	.242	4.011
77°	.974	.225	4.331
78°	.978	.208	4.705
79°	.982	.191	5.145
80°	.985	.174	5.671
81°	.988	.156	6.314
82°	.990	.139	7.115
83°	.993	.122	8.144
84°	.995	.105	9.514
85°	.996	.087	11.43
86°	.998	.070	14.30
87°	.999	.052	19.08
88°	.999	.035	28.64
89°	1.000	.017	57.29
90°	1.000	.000	

Answers to Odd-numbered Questions

Chapter 1

1. 1.85 m.

3. 0.00782 cm $= 7.82 \times 10^{-3}$ cm.

5. 0.276 in.; $0.0230 = 2.30 \times 10^{-2}$ ft.

7. 4.31×10^{4}.

9. (a) 4×10^{-5}; (b) 0.523.

11. 1.10×10^{27}.

13. 8.5 lb; 4 lb.

15. $10^{-3} \times \sqrt{10} = 3.16 \times 10^{-3}$ m $= 3.16$ mm.

17. 10^{12}.

19. (a) 2.7×10^{-17}; (b) 2.12×10^{5}; (c) 9.73×10^{13}.

21. (a) 10^{3}; (b) 2×10^{3}; (c) 3.16×10^{3}; (d) 3.16×10^{3}; (e) 7×10^{3}; (f) 5.62×10^{5}.

23. (a) 10^{2}; (b) 2×10^{2}; (c) 2.16×10^{2}; (d) 2.16×10^{2}; (e) 6×10^{2}; (f) 6.80×10^{2}.

Chapter 2

1. 35 lb.

3. (a) Yes; (b) 167 lb.

5. $\sum \tau = 0$.

7. 32 lb and 16 lb.

9. 2.29 ft from heavy end.

11. (a) 14 lb; (b) 1.29 ft from 8-lb end; (c) 9.33 lb and 4.67 lb.

15. (a) 17.32; (b) 10.

17. (b) 3.41 N; 12.07 E.

19. 12.54, 15.8° N of E.

21. 7300 lb, 7.5° S of E.

23. (a) 199.4 lb; (b) 159.2 lb.

25. (a) 180 lb; (b) 156 lb, 50 lb.

27. 35.4 lb.

29. 15 kg downward and to the right, 36.9° from vertical; applied 2 m from right end.

31. 0.21.

35. 51.3°.

37. 55.2 lb.

39. 62.4 lb/ft^3.

41. 1780 gm or 1.78 kg.

43. (a) 430 gm; (b) 394 gm; (c) 2950 gm.

45. 4.33 lb/in^2.

47. 1.50×10^5 lb.

49. 462.5 gm/cm^2.

51. 13.9 lb.

53. (a) 100 gm; (b) 100 gm; (c) 100 gm; (d) 100 cm^3; (e) 100 cm^3; (f) 2.50 gm/cm^3.

55. (a) 500 cm^3; (b) 6.00 gm/cm^3.

57. (a) 100 cm^3; (b) 2.50 gm/cm^3; (c) 1.25 gm/cm^3.

59. 0.60.

61. Level drops.

63. (a) 1760 kg; (b) 2030 kg.

Chapter 3

1. (a) 6.0 mi/hr/sec or 8.8 ft/sec^2; (b) 440 ft.

3. 10.1 ft/sec^2.

5. (a) 5×10^{-11} sec; (b) 3.75×10^{-4} cm.

7. (a) 0.0432 m/sec^2; (b) 35.6 m/sec.

9. (a) 0.25 m/sec^2; (b) 1.00 m/sec.

13. 5400 dynes; 5.40×10^5 dynes or 5.40 nt.

15. 0.0625 slug.

17. (a) 44.4 cm/sec^2; (b) 1.33×10^4 dynes.
19. (a) 44.4 cm/sec^2; (b) 7.21×10^4 dynes.
21. (a) 0.492 ft/sec^2; (b) 340 ft.
23. 2.74 sec.
25. 256 ft.
27. (a) 12 m/sec^2; (b) 60 nt.
29. (a) 9.70 cm/sec^2; (b) 6.44 sec.
31. 9.897×10^5 dynes.
33. 108.9 cm/sec^2.
35. 21.8 cm/sec.
37. 2.40×10^5 dynes.
39. 202.5 lb.
41. (b) 2.45 m/sec^2.
43. 10.9 ft/sec^2 or 3.35 m/sec^2 or 335 cm/sec^2.
45. 4.93 ft/sec^2 or 1.51 m/sec^2 or 151 cm/sec^2.
47. 31.0 lb.
51. 20.9 ft/sec.
53. 3.43 sec.
55. 10^5 dynes.
57. (a) 81 ft; (b) 432 ft.
59. (a) 74.5 ft/sec; (b) 65.0 ft.
61. (a) 8.55 sec; (b) 1111 m.

Chapter 4

1. 1.98×10^5 ft-lb.
3. (a) 7200 ft-lb; (b) 7200 ft-lb.
5. (a) 10-lb weight (80 vs. 60); (b) 20-lb weight (260 vs. 180).
7. 2.57×10^5 joules.
9. 2×10^{-8} erg or 2×10^{-15} joule.
11. 9.
13. 7.00 m/sec.
15. 6×10^6 ergs.
17. 1.024×10^5 ft-lb.
19. 13.9 lb.
21. 76.7 cm/sec.
23. 1350 dynes/cm^2.
25. 866 cm^3/sec.
27. 144 ft-lb/sec or 0.26 horsepower.
29. 15.2 horsepower.
31. 1633 watts; 2722 watts.
33. 625 lb.

35. 1.44×10^6 joules.

37. 7.00 cm/sec.

39. 2.5×10^4 sec.

41. 4.52×10^7 cm/sec.

43. 37.5 mi/hr.

45. $v_p = 6 \times 10^4$ cm/sec west; $v_\alpha = 4 \times 10^4$ cm/sec east.

47. (a) 0.599 m/sec; (b) 0.998.

49. 4 kg/sec.

51. 76.6 lb.

53. 306 lb.

Chapter 5

1. 0.52 rad.

3. (a) 0.262 rad/hr; (b) 1047 mi/hr.

5. (a) 251 rad/sec; (b) 83.8 ft/sec.

7. 62.8 rad/sec^2.

9. 131 rev.

11. 9×10^5 gm-cm^2 or 0.090 kg-m^2.

13. 2.67 kg-m^2.

15. (a) 1.5×10^5 gm-cm^2 or 0.015 kg-m^2; (b) 1.80×10^7 dyne-cm or 1.80 nt-m.

17. 0.0208 slug-ft^2.

19. 38.4 lb.

21. 17,020 lb.

23. (a) 3.57×10^{27} dynes or 3.57×10^{22} nt; (b) 3.64×10^{18} metric tons.

25. (a) 140 cm/sec; (b) 9800*m* ergs; (c) 39,200*m* ergs; (d) 49,000*m* ergs; (e) 0; (f) 49,000*m* ergs; (g) 313 cm/sec.

27. Will stay in place. $F_f = 0.98m$; $F_c = 0.79m$.

29. 1.42 joules.

31. $\sqrt{2gh}$.

33. $\frac{1}{2}$ transl., $\frac{1}{2}$ rotat.

35. $\frac{1}{2}$ rev/sec.

37. Ratio of KE's is 2: 1.

39. (a) 120 rev/min; (b) the same in either case; (c) KE before $= 10^8 \times \pi^2$ ergs, or 98.70 joules, KE after $= 2 \times 10^7 \times \pi^2$ ergs, or 19.74 joules; ratio = 5: 1; (d) 1.26 nt-m.

41. Nose down.

Chapter 6

1. 0.107 dyne.

3. 1.08 cm/sec^2.

5. dyne-cm^2/gm^2 or cm^3/(gm-sec^2).

7. 10:8.
9. (a) 3.12×10^{11} joules; (b) half as much.
11. (a) 6.24×10^{11} ergs; (b) -6.24×10^{11} ergs.
13. (a) 15.8 km/sec; (b) 7.92 km/sec.
15. 14.9 km/sec or 9.28 mi/sec.
17. 10.5 km/sec or 6.56 mi/sec.
19. 4.40 mi/sec or 7.08 km/sec.
21. 23 mi/sec.
23. 2.76 Martian radii.
25. 5.27 Martian radii.

Chapter 7

1. 500 lb/in^2.
3. 9.98×10^9 dynes/cm^2.
5. 1.67×10^{-4}.
7. 2×10^{12} dynes/cm^2.
9. 7.85 lb.
11. (a) 1/3; (b) 1/4.
13. 5×10^{-4} sec.
15. (a) 2.5 cycles/sec; (b) 0.4 sec; (c) 2.5 rev/sec; (d) 15.7 rad/sec.
19. (a) 157 cm/sec; (b) 4935 cm/sec^2.
21. 134 vibr/min.
23. 94.7 vibr/min.
25. 132 cycles/min.
27. 99.3 cm.
29. 2.452 hr.
31. 0.027 radian or 1.5 degree.
33. No. Natural frequency is 299.6 vibr/min.

Chapter 8

3. 110 m/sec.
7. 0.78 sec.
9. At $t = 0.78$ sec and 1.56 sec.
11. 5×10^{14} waves/sec.
13. 1100 ft/sec.
15. 90 ft/min.
17. (a) 8 in.; (b) 1650/sec.
19. 6.55×10^8 dynes.

Chapter 9

1. 4400 ft.
3. 0.73 sec.

5. 1080 ft/sec.
7. 0.020 sec.
9. 2160 vibr/sec.
11. 440 vibr/sec.
13. 208 vibr/sec.
15. (a) 550/sec; (b) 1100/sec; (c) 1650/sec.
17. 880 vibr/sec; 1760 vibr/sec.
19. 1.25 ft or 37.5 cm.
21. 2500.
23. 15 knots.
25. 40.3°.
27. 100 ft/sec.
29. (a) 2128 vibr/sec; (b) 2120 vibr/sec; (c) 1926 vibr/sec.

Chapter 10

1. 37.0°C.
3. 40°C, 104°F.
5. 170°F.
7. (a) 393°K; (b) 1089.1°K; (c) 243°K; (d) 211.9°K.
9. 14.7 lb/in^2.
11. 12.5 atmospheres or 184 lb/in^2.
13. (a) 22.5 lb/in^2; (b) 7.8 lb/in^2.
15. 36.5 lb/in^2 if atm. pr. = 15 lb/in^2; 35.5 lb/in^2 if atm. pr. = 10 lb/in^2.
17. 3940 ft^3.
19. 299.90 ft.
21. (b) 211°C.
23. 1006 cm^3.
25. 100.32 cm.
27. 924 cal.
29. 9.62×10^5 cal.
31. 63.4°C.
33. 63.5°C.
35. (a) 80 gm; (b) 38.7 gm.
37. 47.2°C.
39. 173 cal/sec.
41. 29,200 sec.
43. 1.49: 1.

Chapter 11

3. 1.76 cal.
5. (a) 4.5×10^9 ergs or 450 joules; (b) 108 cal.

7. (a) 1876 cal; (b) 6.25°C.
9. 106°C.
11. 0.00732.
13. 364 sec.
15. 0.505.
17. 0.167.
19. (a) −214 cal/deg; (b) +293 cal/deg; (c) +79 cal/deg.

Chapter 12

1. 4.8 dynes attraction.
3. 13.9 esu.
5. 2.77×10^{-7} cm.
7. (b) both +12 esu or both −12 esu; (c) −12 esu and +36 esu or +12 esu and −36 esu.
15. 1.2 dynes/esu, north.
17. (a) 1.07×10^{4} dynes/esu; (b) 5.12×10^{-6} dyne.
19. 4 cm from the 64-esu charge, toward the 256-esu charge.
21. 5.87×10^{-6} dyne.
23. +12 ergs/esu.
25. 15.36×10^{-13} erg.
27. (a) No; (b) 2.4 cm from −64 esu charge toward 256 esu charge, and 4.0 cm from −64 esu charge away from 256 esu charge.
29. (a) 0 volt; (b) −36,800 volts; (c) 36,800 volts; (d) 0.0368 joule.
31. 22*C* to 26*C*, depending on type of mica.
33. (a) 10^{4} volts/m or 10^{4} nt/coul; (b) 2×10^{4} volts/m or 2×10^{4} nt/coul.
35. 50 volts.
37. Increased.
39. 45.5 volts.
41. 6.64×10^{-4} μf.

Chapter 13

1. 30 ohms.
3. 0.6 volt.
5. 1.56 ohm.
7. 2.4 amp.
9. 0.4 amp.
11. (a) 15 ohms; (b) 0.4 amp; (c) 0.6 volt; (d) 5.4 volts; (e) 1.8 volts.
13. 5.5 ohms.
15. 1 ohm.
17. 2.00 amp, 2.67 amp, 4.00 amp, 8.67 amp.
19. (a) $I_3 = 1.00$ amp, $I_5 = 1.5$ amp, $I_6 = 0.50$ amp, $I_B = 1.5$ amp; (b) 10.5 volts; (c) 7.5 volts; (d) 3.0 volts.

21. (a) 3 watts; (b) 33.3 sec.
23. 4.47 volts.
25. (a) 360 ohms; (b) 22.5 watts.
27. (a) 22.8 ohms, in series; (b) 0.95.
29. 25-watt lamp.
31. 8020 sec.

Chapter 14

1. 47.7 amp/m.
3. 938 amp/m.
5. 12.6 amp.
7. 6.00×10^{-5} weber/m^2.
9. (a) 3.2×10^{-17} nt; (b) away from wire.
11. Upward.
13. (a) 3.84×10^{-14} nt; (b) 5.77×10^{12} m/sec.^{2}
15. (c) 6.24 cm.
17. (a) 1.60×10^{-17} joule; (b) 5.93×10^6 m/sec; (c) 6.75×10^{-4} weber/m^2.
19. 3.77×10^{-5} weber.
21. 3.54×10^{-5} weber.
23. 6000 amp/m, 7.54×10^{-3} weber/m^2.
25. 0.603 weber/m^2.
27. Not much: $F = 2 \times 10^{-5}$ nt.
29. 0.050 weber/m^2.
31. 6.25×10^{11} electrons/sec.
33. 4900 ohms.
35. 0.05002 ohm.
37. (a) Each force 10^{-5} nt = 1 dyne.
39. 5880 amp.
41. 4×10^{-4} weber.
43. (a) 0.1 volt; (b) no.
45. (a) 6×10^{-3} volt; (b) 1.5×10^{-4} amp.
47. 4.46 weber/m^2 or 44,600 gauss.
49. (a) 33,900 volts; (b) 200 amp; (c) 283 amp.
51. 40 volts, 40 amp.
53. \$348 − \$14 = \$334.

Chapter 15

1. 3 ft.
5. (a) 8.00 cm RI; (b) 9.60 cm RI; (c) 16.0 cm RI; (d) 24.0 cm RI; (e) ∞, no image; (f) −8.00 cm VE; (g) −2.67 cm VE.

7. (a) 36 in. behind mirror; (b) virtual; (c) $4\frac{1}{2}$ in.; (d) erect.

9. 6 in.

11. 32.1°.

13. 48.6°.

15. (a) 23.7°; (b) 40°.

17. 52.4°.

19. 33.2°.

21. 27.9°.

23. (a) 8.00 cm RI; (b) 9.60 cm RI; (c) 16.0 RI; (d) 24.0 cm RI; (e) ∞, no image; (f) -8.00 cm VE; (g) -2.67 cm VE.

25. 0.62 mm.

27. 0.476 ft or 5.71 in.

29. 15 cm beyond diverging lens.

31. 15 cm.

33. 6.00 in. in front of screen or 4.65 in. from object, between object and first lens.

35. (a) 1.11 in.; (b) 2.78 cm.

37. 484.

39. 96.

41. 0.90 in., 0.45 in., 0.30 in.

43. $53°58' - 53°12' = 0°46'$.

Chapter 16

3. 500 sec.

5. 2.94×10^8 m/sec.

7. 14.8 ft-candles.

9. 2.41×10^{27} candlepower.

11. 320 candlepower.

13. cp/watt (40) $= 1.50 \times$ cp/watt (10).

23. 1.99×10^{10} cm/sec.

Chapter 17

1. (a) 1.667 cm; (b) 11; (c) 12.

5. 0.38 mm.

7. 2.55×10^{-5} cm or 2550 Å; not visible.

9. (a) 18.1°; (b) 38.3°; (c) 68.4°.

11. 2.435×10^{-5} in., or 6190 Å.

13. (a) Second order.

19. (a) $\lambda_0 = 3.6188 \times 10^{-5}$ cm, $\lambda_E = 4.0377 \times 10^{-5}$ cm; (b) $N_0 = 2763.3$, $N_E = 2476.6$; (c) 286.7 wavelengths.

21. 4×10^{-5} cm or 4000 Å.

Chapter 18

1. (a) 0.1680 hr; (b) 0.1693 hr.

3. (a) $(6.67 \times 10^{-8} + 3.33 \times 10^{-16})$ sec; (b) $(6.67 \times 10^{-8} + 6.67 \times 10^{-16})$ sec.

5. $0.8c$ or 2.4×10^8 m/sec.

7. $-c$ or -3×10^8 m/sec.

9. 80 m.

11. 0.25 mm.

13. 6.46×10^{-27} gm.

15. 5×10^{-9}, 3×10^{13} metric tons.

17. 2.60×10^{10} cm/sec.

19. 1.33 min.

21. 60 years old.

25. 0.35 gm.

27. 3×10^{19} gm; yes.

Chapter 20

1. 7.8 diameters.

3. 0.021 cm/sec.

5. (a) 127°C; (b) 527°C.

7. 471 m/sec.

9. (a) Same; (b) $v_B = 0.707\, v_A$; (c) same.

11. Rms = 3.89, av = 3.50.

13. v_{rms} (hydrogen) = $1.41 v_{rms}$ (helium).

15. 1.31×10^5 cm/sec.

17. (a) More; (b) more; (c) increased.

Chapter 21

3. (b) $m_{Fe} = 1.43 m_K$; (c) 1.43; (d) same.

5. Number Fe atoms = $0.293 \times$ number Zn atoms.

7. (a) 1.008 gm; (b) 8.00 gm; (c) 31.77 gm; (d) 8.99 gm.

9. (a) 1.24 gm Au; (b) 1820 coul.

11. 0.084 cent.

13. 3.2×10^{-15} nt.

15. 2×10^{-3} weber/m².

17. 2560 volts.

19. 9.6×10^{-15} nt.

21. (a) 3.64×10^{-5} erg; (b) 3.64×10^{-5} erg.

23. About 3000 in. or 250 ft.

Chapter 22

1. (a) 9.66×10^{-4} cm or 96,600 Å; (b) 4.83×10^{-4} cm or 48,300 Å.

3. 14,500°K.

5. (a) 459 watts/m^2; (b) 5.60×10^8 watts/m^2.
7. 1780°K.
9. (a) 5.76×10^{-21} erg; (b) 3.62×10^{-12} erg; (c) 3.31×10^{-8} erg.
11. 1.99×10^{-4} cm.
13. 7.49×10^{-12} erg.
15. (a) 2.80×10^{-12} erg; (b) 3.37×10^{-12} erg; (c) 0.57×10^{-12} erg; (d) 3.54×10^7 cm/sec.
17. 1.816×10^{15}/sec.
19. 1.02 Å.

Chapter 23

1. 2.92×10^{15}/sec; 3.16×10^{15}/sec.
3. $F_G/F_E = 4.42 \times 10^{-40}$.
5. 4.77 Å.
7. 0.27 Å.
9. -2.18×10^{-11} erg.
11. (a) 8.72×10^{-11} erg; (b) -17.44×10^{-11} erg.
13. (a) 1.216×10^{-5} cm or 1216 Å; (b) no—ultraviolet; (c) Lyman.
15. Each line has $f = 4 \times$ frequency of analogous hydrogen line.

Chapter 24

1. (a) 0, 1, 2; (b) $-3, -2, -1, 0, 1, 2, 3$.
5. (a) 2.2×10^{-23} cm^3; (b) 1.4×10^{-23} cm^3.
7. 0, 1, 2, 3.
9. $n = 1, l = 0, m = 0, s = \frac{1}{2}$; $n = 1, l = 0, m = 0, s = -\frac{1}{2}$; $n = 2, l = 0, m = 0, s = \frac{1}{2}$; $n = 2, l = 0, m = 0, s = -\frac{1}{2}$.
11. $n = 1, l = 0$: 2; $n = 2, l = 0$: 2; $n = 2, l = 1$: 6; $n = 3, l = 0$: 2; $n = 3, l = 1$: 6; $n = 4, l = 0$: 1.
13. $l = 3, n = 4$.
15. (a) 3; (b) 2.
19. 0.994 Å.

Chapter 25

1. 3.64×10^{-7} cm.
3. 2.27×10^8 cm/sec.
5. 6.66×10^{-8} cm.
7. (a) 2.35×10^{-11} erg or 2.35×10^{-18} joule; (b) 14.7 volts.
9. (a) $v_e = 42.9 v_p$; (b) $M_p = 42.9 M_e$; (c) $\lambda_e = 42.9 \lambda_p$.
11. 1.42×10^{-8} cm.
13. 2.48×10^{-9} cm.
15. 1.06×10^{-12} cm/sec.
17. 2.33×10^{-5} cm.
19. (a) 23.2 cm/sec.

Chapter 26

1. C: 6P, 6N; Bi: 83P, 126N; Sc: 21P, 24N; Au: 79P, 118N.
3. 6.91×10^6 cm/sec.
5. 35.5.
7. ${}_8O^{16}$, ${}_8O^{17}$, ${}_8O^{18}$; ${}_{30}Zn^{64}$, ${}_{30}Zn^{66}$, etc.
9. (a) ${}_{87}Fr^{221}$; (b) ${}_{88}Ra^{219}$; (c) ${}_{90}Th^{234}$.
11. (a) ${}_{89}Ac^{225}$; (b) ${}_{87}Fr^{223}$; (c) ${}_{95}Am^{243}$.
13. ${}_{88}Ra^{226} \longrightarrow {}_2He^4 + {}_{86}Rn^{222}$.
15. ${}_{19}K^{40} \longrightarrow \beta + {}_{20}Ca^{40}$.
17. ${}_{19}K^{40} + \beta \longrightarrow {}_{18}Ar^{40}$.
19. (a) 6 α-emissions; (b) 4 β-emissions.
21. (a) 0.0652 Å; (b) 3.05×10^{-7} erg = 0.191 Mev.
23. 5×10^5 ev, 8×10^{-7} erg, 8×10^{-14} joule.
25. $KE(N^{++}) = 2\,KE(e)$.
27. 0.0100 μgm.
29. Somewhat less than 7 half-lives; about 26.6 months.
31. (a) 0.25 gm; (b) no; (c) yes.
33. 11,400 yr.
35. About 70 years ago.

Chapter 27

1. ${}_1H^1$.
3. ${}_4Be^9 + {}_2He^4 \longrightarrow {}_6C^{12} + {}_0n^1$.
5. (a) ${}_{11}Na^{23} + {}_1H^1 \longrightarrow {}_{10}Ne^{20} + {}_2He^4$.
7. (a) ${}_5B^{11} + {}_1H^1 \longrightarrow 3({}_2He^4) + \gamma$; (b) ${}_{13}Al^{27} + {}_1H^2 \longrightarrow {}_{12}Mg^{25} + {}_2He^4$; (c) ${}_7N^{14} + {}_1H^2 \longrightarrow 4({}_2He^4)$; (d) ${}_{13}Al^{27} + {}_1H^2 \longrightarrow {}_{13}Al^{28} + {}_1H^1$; (e) ${}_{13}Al^{27} + {}_1H^1 \longrightarrow \gamma + {}_{14}Si^{28}$.
9. 1.6×10^{-6} erg = 1.6×10^{-13} joule.
11. 37.5 rev.
13. 8.19×10^{-8} sec.
15. 6.10×10^6 cycles/sec.
17. 1.70×10^7 m/sec.
19. 89 cm.
21. 1.16 $\times$ nonrelativistic time = increase of 0.16.
23. (a) 1.33×10^{11} cm/sec; (b) no; (c) it is not impossible.

Chapter 28

1. (a) 1.14; (b) 1.33; (c) 1.52.
3. (a) 1.66×10^{-23} gm; (b) 6.9×10^{-38} cm³; (c) 2.55×10^{-13} cm.
5. 3.1 cm.
7. (a) 0.03040 amu = 28.3 Mev; (b) 28.3 Mev.
9. 7.08 Mev/nucleon.
11. 0.00944 amu = 8.79 Mev/nucleon.

13. 7.57 Mev/nucleon.

15. 0.00612 amu = 5.70 Mev product energy.

17. 0.00201 amu = 1.87 Mev bombarding particle energy.

19. 0.00256 amu = 2.38 Mev product energy.

21. (a) About 50 Mev released; (b) about 12 Mev supplied.

23. (a) About 20 Mev released; (b) about 15 Mev supplied.

Chapter 29

1. (a) Decreased; (b) increased.

5. About 200 Mev.

7. (a) $U^{235}F_6$; (b) 1.004.

9. (a) 0.33; (b) reduced to 4 percent after 8 collisions.

11. (a) 10^{21} ergs; (b) 6.25×10^{26} Mev; (c) 3.1×10^{24}; (d) about 1200 gm.

13. 1670 ev.

15. About 400/yr, or more than 1 per day.

Chapter 30

1. About 10^9 pairs.

3. 4.3 Bev.

5. About 7×10^5 cm, or 7 km.

Chapter 31

1. +29 in.

3. +18.75 cm.

7. 2.83 times as great.

9. (a) 4.5 and 0.23; (b) about 0.14 cal/hr/cm^2 (man) and 0.07 cal/hr/cm^2 (shrew)—the shrew is a little better insulated with fur; (c) inversely as r; (d) 10^3.

Chapter 32

3. 0.162.

5. Increases.

7. 83.6 ft.

9. 750 ft.

Chapter 33

1. (a) 4000 m/sec; (b) 0.36.

3. $v_{rms} = 3.65 \times 10^6$ cm/sec, $v_{esc} = 6.13 \times 10^7$ cm/sec; $v_{rms}/v_{esc} = 0.06$.

5. 705 Å.

7. (a) Toward; (b) 375 km/sec.

9. 0.044 Å.

11. 3.6×10^{38} protons/sec.

Index

A

B